

# **DISSERTATION**

**zum Erwerb des Doktorgrades der Naturwissenschaften**

**an der Medizinischen Fakultät der**

**Ludwig-Maximilians-Universität München**

## **Functional and biochemical characterization of the epigenetic inheritance of an artificial centromere**

**Stefanie Monika HERMANN**

**2017**



Aus der Abteilung Genvektoren des  
Helmholtz Zentrum München  
Wissenschaftlicher Geschäftsführer: Prof. Dr. Günther Wess

# **Functional and biochemical characterization of the epigenetic inheritance of an artificial centromere**

## **Dissertation**

zum Erwerb des Doktorgrades der Naturwissenschaften  
an der Medizinischen Fakultät der  
Ludwig-Maximilians-Universität München

vorgelegt von  
**Stefanie Monika HERMANN**

aus  
**Wolfratshausen**

**2017**

**Mit Genehmigung der Medizinischen Fakultät  
der Ludwig-Maximilians-Universität München**

Betreuer: Privat Dozent Dr. Aloys SCHEPERS

Zweitgutachter: Prof. Dr. Gunnar SCHOTTA

Dekan: Prof. Dr. med. dent. Reinhard HICKEL

Tag der mündlichen Prüfung: 20.03.2018

# Eidesstattliche Versicherung

Hermann, Stefanie Monika

---

Name, Vorname

Ich erkläre hiermit an Eides statt,  
dass ich die vorliegende Dissertation mit dem Titel

## **Functional and biochemical characterization of the epigenetic inheritance of an artificial centromere**

selbstständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter der Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

Mering, 26. März 2018



*Für meine liebe Mutter und Christian,  
die mich während dieser Arbeit immer unterstützt haben*





*„Man muss an seine Berufung glauben  
und alles daransetzen, sein Ziel zu erreichen.“*

- Marie Curie -



# TABLE OF CONTENT

<b>TABLE OF FIGURES</b>	<b>V</b>
<b>TABLE OF TABLES</b>	<b>VIII</b>
<b>ABBREVIATIONS</b>	<b>IX</b>
<b>ZUSAMMENFASSUNG</b>	<b>XI</b>
<b>ABSTRACT</b>	<b>XIII</b>
<b>1 INTRODUCTION</b>	<b>1</b>
1.1 Chromatin and histone modifications	2
1.2 Mitosis	4
1.3 The kinetochore	5
1.3.1 Inner kinetochore	6
1.3.2 Outer kinetochore	8
1.4 The Centromere	9
1.4.1 Centromeres in different organisms	10
1.4.2 CENP-A determines centromere identity	11
1.5 The centromeric histone H3 variant CENP-A	12
1.5.1 CENP-A structural domains	13
1.5.2 CENP-A nucleosome	14
1.6 Human centromeric chromatin	15
1.7 <i>De novo</i> centromere formation	18
1.7.1 Human artificial chromosomes (HACs)	19
1.7.2 Centromere formation at ectopic sites on chromosomes	20
1.7.3 <i>De novo</i> Centromere formation on LacO plasmids	21
1.8 Development of the pCON <sup>CENP-A</sup> plasmid system	23
<b>2 AIM OF THE THESIS</b>	<b>29</b>
<b>3 MATERIALS &amp; METHODS</b>	<b>31</b>
3.1 Materials	31
3.1.1 Devices and consumables	31
3.1.2 Software	33

## TABLE OF CONTENT

3.1.3	Enzymes and antibodies	33
3.1.4	Chemicals	35
3.1.5	Kits	37
3.1.6	Buffer	37
3.1.7	Plasmids	39
<b>3.2</b>	<b>Methods</b>	<b>40</b>
3.2.1	Molecular biological methods	40
3.2.1.1	Cloning strategies	40
3.2.1.2	Polymerase Chain Reaction (PCR)	42
3.2.1.3	Generation of chemical competent DH5 $\alpha$ bacteria	43
3.2.1.4	Transformation	43
3.2.1.5	DNA preparation (Mini-prep.)	44
3.2.1.6	Sequencing	44
3.2.1.7	High amount DNA preparation (Maxi-prep.)	45
3.2.1.8	Restriction digest and dephosphorylation	45
3.2.1.9	Agarose gel electrophoresis	46
3.2.1.10	Gel extraction	46
3.2.1.11	Ligation	46
3.2.2	Cell culture	47
3.2.2.1	Cultivation of HEK293 cell lines	47
3.2.2.2	Determination of cell number	48
3.2.2.3	Generation of stable cell lines	49
3.2.2.4	Cryoconservation	49
3.2.2.5	Transfection for plasmid maintenance experiments	49
3.2.2.6	Transient transfection	50
3.2.2.7	Cell synchronization with thymidine block	50
3.2.2.8	Transient transfection of synchronized cells	51
3.2.2.9	Transient transfection for live cell imaging	52
3.2.3	Plasmid rescue assay	52
3.2.4	Flow cytometry	53
3.2.5	Western Blot	54
3.2.6	Co-Immune precipitation of scetR fusion proteins	54
3.2.7	Chromatin Immune Precipitation (ChIP)	55
3.2.7.1	Cross-link	56
3.2.7.2	Sonication	56
3.2.7.3	Immune precipitation	56
3.2.8	Quantitative PCR (qPCR)	57
3.2.9	STREP pull down with Cas9:mCherry:TAP	58
3.2.9.1	Cross-link	59
3.2.9.2	French Press	59
3.2.9.3	Restriction digest	60
3.2.9.4	Pull down	61
3.2.10	Live cell imaging with Cas9:3xmCherry	62
3.2.11	Immune fluorescence	62
3.2.11.1	CENP-C/Ndc80 and EBNA1 co-staining	62
3.2.11.2	EBNA1 staining	63
3.2.12	Microscopy	64

## TABLE OF CONTENT

3.2.12.1	Fixed samples	64
3.2.12.2	Live cell Imaging	64
<b>4</b>	<b>RESULTS</b>	<b>66</b>
<b>4.1</b>	<b>Long term pCON<sup>CENP-A</sup> maintenance</b>	<b>67</b>
<b>4.2</b>	<b>sctetR:CENP-A is site specifically targeted to tetO sites</b>	<b>69</b>
4.2.1	ChIP directly demonstrates interaction with tetO	69
4.2.2	sctetR:GFP:CENP-A co-localizes with plasmids	71
<b>4.3</b>	<b>Dynamics of pCON<sup>CENP-A</sup> during cell cycle</b>	<b>73</b>
4.3.1	CRISPR/Cas9 targeting system	73
4.3.2	Cas9:3xmCherry is suitable to visualize pCON <sup>CENP-A</sup> plasmids	76
4.3.3	pCON <sup>CENP-A</sup> localization in living cells	77
<b>4.4</b>	<b>Recruitment of kinetochore proteins</b>	<b>80</b>
4.4.1	Recruitment of inner kinetochore proteins (CENP-C)	80
4.4.2	Recruitment of outer kinetochore proteins (Ndc80)	82
4.4.3	Recruitment of kinetochore proteins only if CENP-A is targeted	83
<b>4.5</b>	<b>sctetR:CENP-A forms centromeric nucleosomes</b>	<b>85</b>
4.5.1	sctetR:GFP:CENP-A is present at endogenous centromeres	85
4.5.2	sctetR:CENP-A nucleosomes	86
<b>4.6</b>	<b>Establishment of centromere inheritance</b>	<b>88</b>
4.6.1	Targeting of sctetR:CENP-A after 5 days but not after 3 weeks	89
4.6.2	Centromere inheritance is established within 7 days	90
4.6.3	Establishment of centromere inheritance after 4 days	93
<b>4.7</b>	<b>Recruitment of CENP-A</b>	<b>96</b>
<b>4.8</b>	<b>CRISPR/Cas9 targeting system to analyze histone modifications on the artificial centromere</b>	<b>99</b>
4.8.1	Specific binding of Cas9 to the targeting sites in the plasmid system	100
4.8.2	High digest efficiency of plasmid DNA	101
4.8.3	Re-Solubilization by high pressure in a French pressure cell press	102
4.8.4	Efficient STREP pull-down of Cas9:mCherry:TAP fusion protein	104
4.8.5	Co-purification of modified histones	106
<b>5</b>	<b>DISCUSSION</b>	<b>108</b>
<b>5.1</b>	<b>Plasmid segregation mechanism after CENP-A targeting</b>	<b>108</b>
<b>5.2</b>	<b>Establishment of self-propagating centromeres on plasmids</b>	<b>112</b>
<b>5.3</b>	<b>Four days are sufficient for <i>de novo</i> Centromere inheritance</b>	<b>114</b>
<b>5.4</b>	<b>Histone modifications at matured artificial centromeres</b>	<b>115</b>

## TABLE OF CONTENT

---

<b>6 CONCLUSION</b>	<b>119</b>
<b>BIBLIOGRAPHY</b>	<b>121</b>
<b>APPENDIX</b>	<b>128</b>
Establishment of a suitable tetR antibody for ChIP and Co-IP	128
tetR antibody related methods	129
tetR antibody validation	129
Covalent coupling of antibodies to protein G sepharose beads	129
Screening of primary antibody supernatants	130
Immune precipitation with primary antibody supernatants	131
ChIP with primary vs. stable generated antibody supernatant	133
Targeting independent plasmid maintenance after 4 days	136
<b>ACKNOWLEDGEMENT</b>	<b>137</b>

TABLE OF FIGURES

Figure 1: Nucleosome core particle.....	2
Figure 2: Mitotic chromosome segregation.....	4
Figure 3 Ultrastructure of vertebrate kinetochore .....	6
Figure 4: Model for human kinetochore assembly .....	8
Figure 5: Schematic representation of different centromeres.....	11
Figure 6: Primary sequence and secondary structure of CENP-A .....	13
Figure 7: Structural model of CENP-A and H3 nucleosome.....	15
Figure 8: Model of epigenetic modifications on centromeric chromatin.....	17
Figure 9: Schematic representation of CENP-A (CID) targeting to LacO sites in chromosomes .....	20
Figure 10: Schematic representation of CENP-A (CID) targeting to LacO plasmids .....	21
Figure 11: Microtubules attach to plasmids.....	22
Figure 12: HAC and <i>oriP</i> maintenance during cell division.....	24
Figure 13: <i>oriP</i> plasmid based on latent replication origin of EBV .....	25
Figure 14: tetO-DS plasmid maintenance .....	26
Figure 15: Plasmid maintenance of different reporter plasmids .....	27
Figure 16: All-in-one pCON <sup>CENP-A</sup> reporter plasmid .....	28
Figure 17 schematic representation of Cas9:mCherry:TAP fusion protein.....	41
Figure 18 Representation of tetO-DS scetR:CENP-A + Cas9 plasmid.....	60
Figure 19: Plasmid rescue and FACS analysis of plasmid maintenance .....	68
Figure 20: ChIP with $\alpha$ -tetR monoclonal antibody .....	70
Figure 21: Illustration of EBNA1 immune fluorescence staining .....	71
Figure 22: Immune fluorescence of EBNA1 and scetR:GFP:CENP-A .....	72
Figure 23: Naturally occurring and engineered CRISPR-Cas9 systems .....	74

## TABLE OF FIGURES

Figure 24: sgRNA targeting sequence and spacer monomer .....	75
Figure 25: sgRNA targeting sequence with restriction sites.....	75
Figure 26: Cas9:3xmCherry is specifically targeted to plasmids in IF .....	76
Figure 27: Highly dynamic localization of pCON <sup>CENP-A</sup> in interphase .....	78
Figure 28: plasmid segregation in mitosis is asymmetric .....	79
Figure 29: Recruitment of the inner kinetochore protein CENP-C.....	81
Figure 30: Recruitment of the outer kinetochore protein Ndc80 .....	83
Figure 31: Quantification of kinetochore protein co-localization after targeting sctetR:GFP:CENP-A or sctetR:GFP .....	84
Figure 32: Immune fluorescence of tubulin .....	85
Figure 33: schematic representation of H3 and CENP-A nucleosomes.....	86
Figure 34: tetR IP and co-IP of endogenous histones .....	87
Figure 35: ChIP of EBNA1 and sctetR:CENP-A in <i>oriP</i> and pCON <sup>CENP-A</sup> plasmid .....	89
Figure 36: Targeting independent plasmid maintenance 7 days after establishment .....	92
Figure 37: Two days of establishment are not enough.....	94
Figure 38: ChIP in presence and absence of doxycycline .....	95
Figure 39: Recruitment of RFP:CENP-A to plasmids .....	96
Figure 40: Recruitment of RFP:CENP-A to plasmids in presence of doxycycline .....	97
Figure 41: ChIP of plasmid with and without targeting sites .....	100
Figure 42: High digest efficiency in un-cross-linked and heavily cross-linked sample .....	101
Figure 43: Protein re-solubilization by French Press.....	103
Figure 44: DNA re-solubilization by French Press .....	104
Figure 45: STREP pull down after French press and digest .....	105
Figure 46: Co-purification of histones and specific modifications .....	106
Figure 47: Segregation mechanism of pCON <sup>CENP-A</sup> plasmids .....	110



## TABLE OF FIGURES

---

Appendix Figure 1: Western Blot of sctetR:HMGA1a <sup>+</sup> cell lysate with different tetR antibodies.....	130
Appendix Figure 2: Immune precipitation with different primary antibody supernatants..	132
Appendix Figure 3: Immune precipitation with 31B3 and 26G3 primary and stable supernatant.....	133
Appendix Figure 4: ChIP with and without doxycycline.....	134
Appendix Figure 5: Targeting independent plasmid maintenance 4 days after establishment .....	136

## **TABLE OF TABLES**

Table 1 Devices and consumables used in this work .....	31
Table 2 Software used during this work .....	33
Table 3 Enzymes used in this work .....	33
Table 4 Antibodies used in ChIP, Co-IP and Western Blot .....	34
Table 5 Antibodies used for immune fluorescence in this work .....	35
Table 6 Substances used in this work .....	35
Table 7 Kits used in this work .....	37
Table 8 Buffer for generation of chemical competent DH5 $\alpha$ .....	37
Table 9 Buffer for plasmid rescue assay .....	38
Table 10 Buffer for cell lysis and Western Blot.....	38
Table 11 Buffer for covalent coupling of antibodies to protein A/G sepharose beads .....	38
Table 12 Buffer for Co-precipitation .....	38
Table 13 Buffer for ChIP and STREP pull down .....	39
Table 14 Reporter plasmids used for ChIP, IF and plasmid maintenance experiments .....	39
Table 15 Plasmids used for CRISPR/Cas9 cloning.....	40
Table 16 Primer for PCR for cloning Cas9:mCherry:TAP.....	41
Table 17 PCR program for cloning Cas9:mCherry:TAP .....	42
Table 18 Essential cell line information .....	48
Table 19 Roche SYBR Green qPCR standard program .....	57
Table 20 qPCR primer .....	58
Table 21 Settings for live cell imaging .....	65

**ABBREVIATIONS**

<b>AT-rich sequence</b>	Nucleic acid sequence containing above average number of adenine and thymine
<b>ATP</b>	Adenosine tri-phosphate
<b>bp</b>	Base pair
<b>Cas9</b>	CRISPR associated protein 9
<b>CATD</b>	CENP-A targeting domain
<b>CCAN</b>	Constitutive centromere associated network
<b>CDEI/II/III</b>	Centromere DNA element I/II/III in budding yeast
<b>CenH3</b>	Centromeric histone H3 variant in plants
<b>CENP-A</b>	Centromere protein A
<b>CENP-C</b>	Centromere protein C
<b>ChIP</b>	Chromatin immuno precipitation
<b>CID</b>	Centromere identifier protein
<b>CIP</b>	Calf-Intestinal-Phosphatase
<b>CLSM</b>	Confocal laser scanning microscopy
<b>CMV</b>	Cytomegalovirus
<b>Co-IP</b>	Co Immuno precipitation
<b>CRISPR</b>	Clustered Regularly Interspaced Short Palindromic Repeats
<b>crRNA</b>	CRISPR targeting RNA
<b>Cse4</b>	Chromosome segregation protein 4
<b>DNA</b>	Deoxyribonucleic acid
<b>dNTP</b>	Deoxynucleotide tri-phosphate
<b>dox</b>	Doxycycline
<b>DS</b>	Dyad symmetry element
<b>EBNA1</b>	Epstein-Barr virus nuclear antigen 1
<b>EBV</b>	Epstein-Barr virus
<b>FACS</b>	Fluorescence activated cell sorting/scanning
<b>FR</b>	Family of repeats
<b>G1</b>	Gap 1 phase during cell cycle
<b>G2</b>	Gap 2 phase during cell cycle
<b>GFP</b>	Green fluorescent protein
<b>H1, H2A, H2B, H3, H4</b>	Histones 1, 2A, 2B, 3 and 4
<b>H3K9ac</b>	Histone 3 lysine 9 acetylation
<b>H3K14ac</b>	Histone 3 lysine 14 acetylation
<b>H4K5ac</b>	Histone 5 lysine 5 acetylation
<b>H4K12ac</b>	Histone 4 lysine 12 acetylation
<b>H3K4me</b>	Histone 3 lysine 4 methylation

## ABBREVIATIONS

---

<b>H3K9me</b>	Histone 3 lysine 9 methylation
<b>H3K27me</b>	Histone 3 lysine 27 methylation
<b>H4K20me</b>	Histone 4 lysine 20 methylation
<b>HA</b>	Hemagglutinin
<b>HAC</b>	Human artificial chromosome
<b>HEC1</b>	Highly expressed in cancer protein
<b>HEK293 cells</b>	Human embryonic kidney 293 cells
<b>HJURP</b>	Holliday junction recognition protein
<b>HMGA1</b>	High mobility group protein 1
<b>HP1</b>	Heterochromatin protein 1
<b>IF</b>	Immune fluorescence
<b>IP</b>	Immuno precipitation
<b>kb</b>	Kilobases
<b>kDa</b>	Kilodalton
<b>KMN</b>	Kn11, Mis12, Ndc80 sub-complex
<b>Kn11</b>	Kinetochores-null protein 1
<b>LacI</b>	<i>Lac</i> repressor
<b>LacO</b>	<i>Lac</i> operator sequence
<b>MNase</b>	Micrococcal nuclease
<b>Ndc80</b>	Nuclear division cycle
<b>NLS</b>	Nuclear localization signal
<b>nt</b>	Nucleotide
<b>OD600</b>	Optical density 600
<b>oriP</b>	Epstein-Barr virus latent replication origin
<b>PAM</b>	Protospacer adjacent motif
<b>PCR</b>	Polymerase chain reaction
<b>PICh</b>	Proteomics of Isolated Chromatin
<b>RFP</b>	Red fluorescent protein
<b>RNA</b>	Ribonucleic acid
<b>RNAi</b>	RNA interference
<b>SAC</b>	Spindle assembly checkpoint
<b>sctetR</b>	Single-chained tetracycline repressor protein
<b>SD</b>	Standard deviation
<b>sgRNA</b>	Single-guide RNA
<b>TAP</b>	Tandem affinity purification
<b>tetO sequence</b>	Tetracycline operator sequence
<b>tetR</b>	Tetracycline repressor protein
<b>tracrRNA</b>	trans-activating crRNA

## ZUSAMMENFASSUNG

Das menschliche Genom unterliegt während der Zellteilung zwei wichtigen Prozessen um seine Integrität zu erhalten. Während der DNA Replikation wird das Genom zunächst vollständig dupliziert. Die duplizierte DNA wird dann verdichtet und während der Mitose gleichmäßig in beide Tochterzellen verteilt. Dies entspricht dem zweiten wesentlichen Prozess des Zellzyklus. Die korrekte Segregation der Schwesterchromatiden wird vorwiegend von Kinetochor-Komplexen reguliert, die sich während der Mitose an den Zentromeren ausbilden. Jedes menschliche Chromosom enthält ein Zentromer an einer definierten Position. Dabei ist die zugrunde liegende DNA Sequenz für die Festlegung der Zentromer-Identität nicht maßgeblich. Vielmehr ist eine zentromer spezifische H3 Histonvariante, das zentromere Protein A (engl. Centromere protein A, kurz CENP-A), das Merkmal für die epigenetische Definition von Zentromeren. Durch künstliches, zielgerichtetes Binden von CENP-A an definierte DNA Sequenzen (engl. „CENP-A targeting“) wird eine vererbare Zentromer-Identität ausgebildet, die sich dann eigenständig erhält. Unser Labor entwickelte ein Plasmidsystem, um die „*de novo*“ Zentromer Entstehung durch „CENP-A targeting“ in humanen Zellen zu untersuchen (pCON<sup>CENP-A</sup>). Diese Plasmide replizieren in Abhängigkeit eines Virusproteins, während zur Segregation „CENP-A targeting“ zu der Ausbildung von Zentromeren führt, wodurch die Plasmide bei der Zellteilung über mehrere Monate stabil erhalten bleiben.

Durch Immunfluoreszenz Analysen von Zellen, die mit pCON<sup>CENP-A</sup> transfiziert wurden, konnte ich zum ersten Mal zeigen, dass die Ausbildung einer Zentromere-Identität durch „CENP-A targeting“ zur Rekrutierung von Komponenten des inneren und äußeren Kinetochors zu den Plasmiden führt. Dadurch wird ein aktiver Segregationsmechanismus induziert, den ich mit Lebend-Zell Mikroskopie untersuchte. Weiterhin konnte ich mit Plasmid-Erhaltungsanalysen nachweisen, dass die Etablierung der vererbaren Zentromere innerhalb von vier Tagen stattfindet. Innerhalb dieses Zeitraumes werden endogene, „targeting“-unabhängige CENP-A

Histone an das Plasmid-Zentromer rekrutiert. Die endogenen CENP-A Proteine wurden in den Immunfluoreszenz Untersuchungen durch RFP:CENP-A Fusionsproteine repräsentiert. Zusätzlich entwickelte ich ein CRISPR/Cas9-abhängiges System um Plasmide mit etabliertem Zentromer aus den Zellen zu reinigen. Mit dieser Methode konnte ich erste epigenetische Histonmodifikationen, die die Vererbbarkeit der Zentromere zusätzlich beeinflussen, nachweisen.

Zusammenfassend untersucht diese Arbeit die „*de novo*“ Etablierung von vererbaren Zentromeren auf Plasmiden in humanen Zellen, die bereits innerhalb von vier Tagen erreicht und durch epigenetische Faktoren bestimmt wird.

## ABSTRACT

The human genome undergoes two major processes during the cell cycle to ensure genome integrity in daughter cell generation. First, the genome is duplicated in a reaction called DNA replication. Duplicated DNA is then compacted and equally distributed to both daughter cells during mitosis. This reflects the second important process during cell division. Major regulator of correct sister-chromatid segregation is the kinetochore complex which is assembled at centromeres. Each human chromosome contains one centromere at a specified position. The underlying DNA sequence at centromeres is not responsible for defining centromere identity. In fact, a centromere specific histone H3 variant, the centromere protein A (CENP-A), is the hallmark for defining centromeres epigenetically. Artificial targeting of CENP-A leads to the formation of an inheritable and self-propagating centromere identity. Our laboratory developed a plasmid system to investigate *de novo* centromere formation in human cells by targeting CENP-A, called pCON<sup>CENP-A</sup>. The replication of these plasmids is dependent on a viral protein, whereas segregation is mediated by the formation of artificial centromeres induced by CENP-A targeting. This leads to a stable maintenance of plasmids in human cells over several months.

By immune fluorescence analysis of cells transfected with pCON<sup>CENP-A</sup>, I showed for the first time that establishment of centromere identity by CENP-A targeting leads to the recruitment of inner and outer kinetochore components. Thus, an active segregation mechanism of the plasmids was induced, that I investigated by live cell imaging microscopy. Furthermore, I demonstrated by plasmid maintenance analysis that establishment of inheritable centromere identity already occurs within four days. Endogenous, targeting-independent CENP-A proteins, represented by RFP:CENP-A fusion proteins, were recruited to plasmids within this timeframe, verified by immune fluorescence experiments. In addition, I developed a CRISPR/Cas9-dependent targeting system for plasmid purification. With this

method I revealed first epigenetic histone modifications, influencing the inheritable centromere identity.

In summary, this work deals with the *de novo* generation of inheritable centromeres on plasmids in human cells, already established after four days and determined by epigenetic factors.



### 1 INTRODUCTION

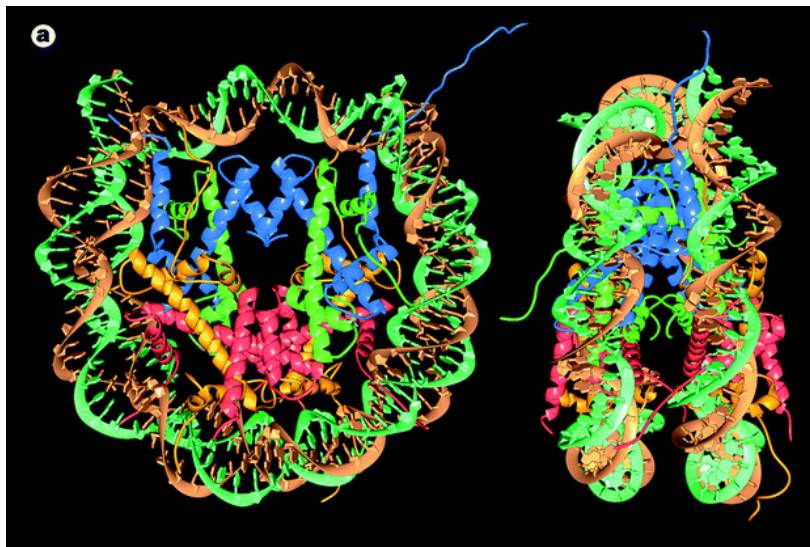
In the 1940s Conrad Waddington introduced the term *epigenetics* and defined it as “the branch of biology which studies the causal interactions between genes and their products which bring the phenotype into being.” (Waddington, 1968). Nowadays, the definition changed more into “the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail a change in DNA sequence” (Wu and Morris, 2001). According to more current literature, epigenetics includes covalent modifications of DNA bases, posttranslational modifications of histones, histone variants and the RNAi pathway (Dupont et al., 2009). Epigenetic modifications fulfil several important functions besides the specification of cells to mitotically inheritable phenotype. They play certain roles in silencing of transposable elements or telomeres, reducing recombination between repetitive elements and ensuring attachment of microtubules to centromeres (Dupont et al., 2009). Centromeres are specific regions on chromosomes mediating kinetochore binding and proper segregation of sister-chromatides during mitosis. Epigenetic marks, like the centromere specific histone variant CENP-A and histone post translational modifications define the centromere position on chromosomes (McKinley and Cheeseman, 2016).

In the following, I will give an overview about chromatin structure, important histone modifications, kinetochore assembly during mitosis and approaches to investigate *de novo* centromere formation.

## 1.1 Chromatin and histone modifications

To fit into the microscopic space of the nucleus, eukaryotic negatively-charged DNA has to be compacted. Therefore, positively-charged proteins, the histones, strongly bind the DNA and pack it into small packing units, the nucleosomes. This chromatin structure is known as a repeating unit of complexes, consisting of four main types of histones and around 200 base pairs, distributed over all DNA since 1974 (Kornberg, 1974). Kornberg characterized these packing units as histone octamers. (Kornberg and Thomas, 1974)

20 years later, in 1997, Luger et al. solved the crystal structure of the nucleosome core particle (Figure 1). (Luger et al., 1997)



**Figure 1: Nucleosome core particle**

Ribbon traces for the 146-bp DNA phosphodiester backbones (brown and turquoise) and eight histone protein main chains (blue: H3; green: H4; yellow: H2A; red: H2B.)

From Luger et al., *Nature* (1997) (Luger et al., 1997)

The nucleosome core particle consists of two of each histones H2A, H2B, H3 and H4 and 146 bp DNA. Histones form an octamer composed of a two H2A/H2B dimers and one (H3)<sub>2</sub>(H4)<sub>2</sub> tetramer and the DNA is wrapped around (Figure 1). Together with the linker Histone H1 and additional linker DNA, the nucleosomes build a

bead-on-a-string structure that can be observed by electron microscopy (Olins and Olins, 1974). Amino-terminal histone tails stick out of the nucleosome core particle. These are prone to be modified and influence the chromatin structure. Different histone-modifying enzymes set modifications mainly at serine, lysine and arginine residues of the N-terminal tails. The best characterized modifications are acetylations and methylations of lysines on H3 and H4 (Dupont et al., 2009). For example, transcriptional activation is often characterized by acetylation of lysine residues on H3 and H4, like H3K9ac, H3K14ac, H4K5ac or H4K12ac. Hyperacetylation of histones leads the formation of euchromatin and DNA is easier accessible due to loosening of DNA packaging (Struhl, 1998). Histone methylation is either activating (H3K4, H3K36, H3K79) or repressive (H3K9, H3K27, H4K20), depending on the amino acid residue and the extend of methylation (mono-, di-, or trimethylation) (Sims et al., 2003).

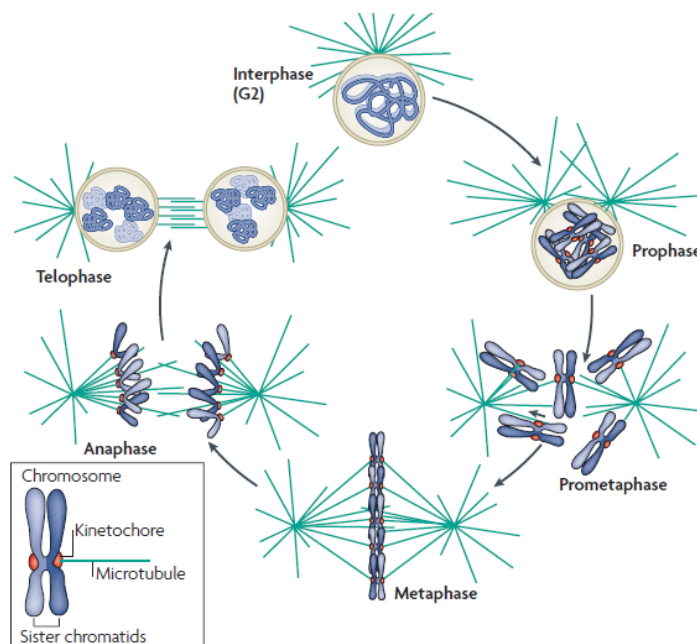
Not only histone modifications determine the chromatin structure. Also histone variants, having some sequence and structural differences from canonical histones, alter the chromatin organization. Histone variants are mainly known for H2A and H3 (Yuan and Zhu, 2012). For example H2A.Z, which is a highly conserved histone H2A variant, differs from canonical H2A in its “docking” domain (Yuan and Zhu, 2012). H3 has, in mammals, three different ubiquitously expressed variants, H3.1, H3.2 and H3.3. H3.1 is only expressed and deposited during S-phase, whereas H3.3 incorporation is replication-independent. Consequently, the H3.3 histone variant is suggested to mediate epigenetic inheritance (Yuan and Zhu, 2012).

The most prominent H3 variant in mammals is the centromere specific H3 variant, centromere protein A (CENP-A). This H3 variant is conserved in nearly all eukaryotes, and the sequence of CENP-A differs significantly from the other histone H3 variants (Henikoff et al., 2004). CENP-A nucleosomes define the structure of centromeric chromatin and are important for regulating mitotic progression.

## 1.2 Mitosis

The eukaryotic cell cycle is separated into four phases that are essential to ensure correct genome maintenance. First, cells duplicate their genetic information during S-phase, a process called DNA replication. Within a gap time, G<sub>2</sub>, the cells prepare for cell division in mitosis. During mitosis replicated and compacted DNA is segregated and equally distributed to the two daughter cells. After exit from mitosis, a second gap time, G<sub>1</sub>, follows before the next replication cycle, or cells rest in G<sub>0</sub> without further proliferation.

Mitosis is separated into five phases: prophase, prometaphase, metaphase, anaphase, and telophase as depicted in Figure 2. Main steps are the breakdown of nuclear envelope and kinetochore assembly in prophase, alignment of sister-chromatids at the metaphase plate and the movement of the chromatids to opposite spindle poles during anaphase (Cheeseman and Desai, 2008).



**Figure 2: Mitotic chromosome segregation**

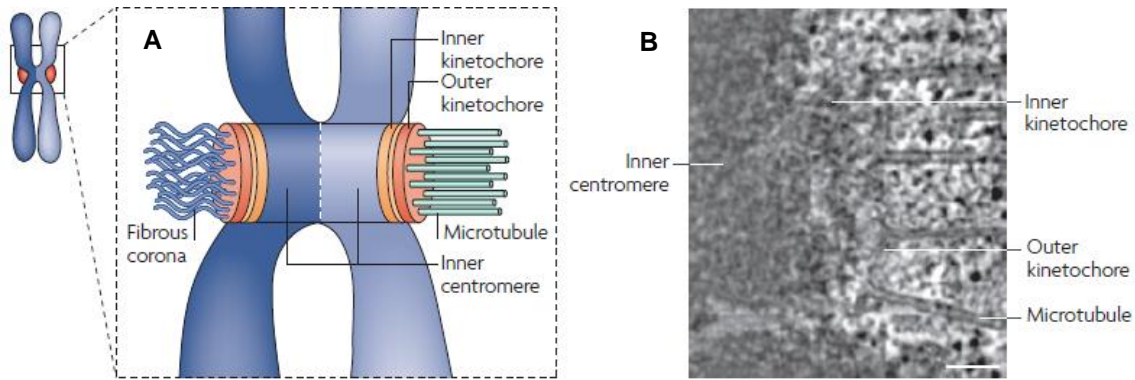
Overview of chromosome-spindle interactions during mitosis.: Assembly of kinetochores on the centromere regions of chromosomes; nuclear envelope breakdown; interaction of kinetochores with spindle microtubules; alignment of chromosomes on the metaphase plate; movement of sister-chromatids move to opposite spindle poles; de-condensation and re-establishment of nuclear envelope  
 Adopted from Cheeseman & Desai, *Nature Reviews, Molecular Cell Biology* (2008)  
 (Cheeseman and Desai, 2008)

The movement to the poles only occurs if sister-chromatids are accurately aligned and each is connected correctly to spindle microtubules. Therefore, the propagation from metaphase to anaphase is tightly regulated by the spindle assembly checkpoint (SAC) (Foley and Kapoor, 2013). SAC is regulated by a huge protein complex that is assembled at the centromeres of chromosomes, called the kinetochore complex. The kinetochore provides an adaptor function between centromeric chromatin on the one hand and microtubules on the other hand. This complex fulfils three main functions. First, it is responsible to maintain cohesion between sister-chromatins at the SAC. Second, it binds microtubules for the correct alignment during metaphase. Third, it monitors microtubule attachment status and signals errors to arrest mitotic progression.

The structure and composition of kinetochores is detailed in the following.

### 1.3 The kinetochore

In an electron micrograph of a vertebrate kinetochore (Figure 3 B), McEwen et al. revealed a kinetochore structure composed of three layers. The inner, electron dense region, is directly associated to the centromere, the outer, also electron dense region, is attached to microtubules. A lighter less electron dense layer separates both regions (McEwen et al., 2007). If kinetochores are not attached to microtubules, for example if drugs are present preventing microtubule polymerization, another fibrous structure, the fibrous corona, is established (Figure 3 A left).



**Figure 3 Ultrastructure of vertebrate kinetochore**

A) Schematic representation of paired sister-chromatins. On the left side the kinetochore is not attached to microtubules and the fibrous corona is established. On the right side the inner and outer kinetochore are represented and the outer kinetochore is attached to microtubules.

B) Electron micrograph of a human kinetochore from McEwen et al. (McEwen et al., 2007)

Image was adopted from Cheeseman & Desai, *Nature Reviews, Molecular Cell Biology* (2008) (Cheeseman and Desai, 2008)

Since the early 2000s kinetochores are examined intensely by mass-spectrometry proteomics and functional genomics. This led to the identification of multiple kinetochore sub-complexes. After purification and reconstitution as well as structural characterization, the structure, composition and stoichiometry of kinetochores becomes clearer (Musacchio and Desai, 2017).

Detailed information about inner and outer kinetochore composition are given in the following chapters.

### 1.3.1 Inner kinetochore

The inner kinetochore directly interacts with the centromeric chromatin and provides a binding platform for the outer kinetochore during mitosis. Inner kinetochore proteins are localized at the centromere throughout the cell cycle and are therefore named the constitutive centromere associated network (CCAN). The CCAN is subdivided into four subunits and CENP-C. The subunits are CENP-LN, CENP-HIKM, CENP-OPQUR and CENP-TWSX (see Figure 4, green boxes).

Only two CENP proteins, CENP-C and CENP-N, of CCAN interact directly and with a higher specificity to CENP-A than to H3 nucleosome (Carroll et al., 2010). CENP-N binds to the CENP-A targeting domain (CATD) (Carroll et al., 2009), whereas CENP-C interacts with the acidic patch of H2A and H2B and the C-terminal tail of CENP-A (Musacchio and Desai, 2017). Binding of CENP-C to the CENP-A tail is dependent on hydrophobic interactions rather than a specific amino acid sequence of CENP-A (Kato et al., 2013). Since CENP-N and CENP-C also bind to canonical H3 nucleosomes, the selectivity is discussed to be enhanced by other CCAN components or posttranslational histone modifications (Bailey et al., 2013; Carroll et al., 2010; Hori et al., 2014).

Another important factor within the CCAN interacting directly with centromeric chromatin is the CENP-B protein. It is the only centromere protein in mammals that recognizes a specific DNA sequence, the CENP-B box. This CENP-B box is a conserved 17 bp sequence and many copies are included in the  $\alpha$ -satellite repeats (Masumoto et al., 1989). CENP-B is not an evolutionary conserved protein, therefore the role of CENP-B in stabilizing centromeres has been questioned. However, CENP-B has some important functions in humans. First, it is important for *de novo* centromere formation in human artificial chromosomes (HAC) (Okada et al., 2007). Second, it contributes to CENP-A nucleosome phasing, CENP-A stabilization and the unwrapping of DNA from CENP-A nucleosomes (Hasson et al., 2013). And third, CENP-B directly binds CENP-C, providing an alternative pathway of CENP-C recruitment (Hoffmann et al., 2016).

The Sub-complexes CENP-LN, CENP-HIKM, CENP-OPQUR and CENP-TWSX contribute to inner kinetochore structure and binding to outer kinetochore proteins during mitosis (Musacchio and Desai, 2017).





The Mis12 complex forms the connection between KMN and inner kinetochore by directly interacting with CENP-C and CENP-T. Mis12 then provides direct binding sites for Knl1 and Ndc80 complexes. The interactions of Mis12 with the CCAN are phosphorylation dependent and therefore regulated in a cell cycle dependent manner by Aurora B kinase (Musacchio and Desai, 2017).

Knl1 is the largest outer kinetochore subunit, has a disordered structure with an array of protein docking sites and is involved in SAC signaling (Musacchio and Desai, 2017).

The best studied sub-complex of KMN is the Ndc80 complex. It is a dumbbell-shaped protein complex and responsible for the microtubule-binding of KMN by the HEC1 subunit (Wei et al., 2007).

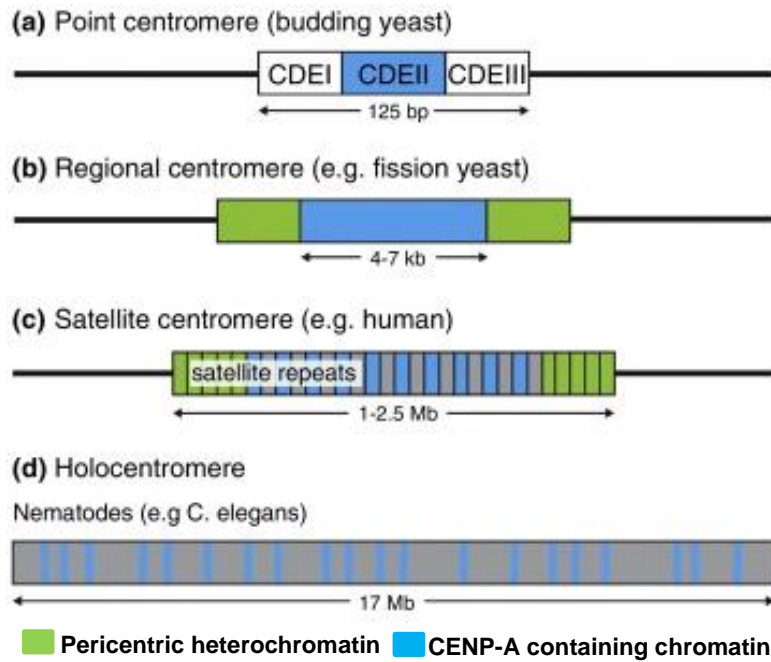
Stoichiometry of kinetochore complex is dictated by linking mechanisms of inner and outer kinetochore proteins as represented in Figure 4. At human centromeres around 20 microtubules are associated with ~100 CENP-A nucleosomes that are dispersed with H3 nucleosomes (Bodor et al., 2014). The centromeric chromatin provides a binding platform for the kinetochore during mitosis. Centromere loss, malfunction and gain of extra centromeres will lead to genome instability or aneuploidy. Hence, it is important to study centromere function in detail and to understand *de novo* centromere formation and inheritance of centromere identity over generations.

### 1.4 The Centromere

Described in the previous chapter, interactions of CCAN with centromeric chromatin are evolutionary highly conserved and, besides CENP-B, not dependent on the underlying DNA sequence. This indicates that centromere identity is epigenetically determined in eukaryotes. In the following I delineate differences in centromere structures in different organisms and specify what these centromeres have in common.

### 1.4.1 Centromeres in different organisms

Since centromeres are of important function and all eukaryotes require centromeres, it is remarkable that there is a large variety of strategies in centromere formation. The smallest and also simplest centromere is found in budding yeast *Saccharomyces cerevisiae*. Centromeric DNA is separated into three domains, CDEI, CDEII and CDEIII, forming a ~120 nt long sequence. Centromere identity is established on CDEII, an AT-rich sequence where one centromeric nucleosome is assembled. This centromere is therefore called a point centromere (Figure 5 a). Fission yeast *Schizosaccharomyces pombe* and other higher eukaryotes establish a larger, regional centromere (Figure 5 b+c). These regional centromeres range from ~5 kb in yeast to megabases in humans and incorporate many centromeric CENP-A nucleosomes. Regional centromeres are flanked by pericentric heterochromatin (Bernard et al., 2001). In humans the centromere is also named as satellite centromere, since the underlying DNA sequence consists of arrays of 171 bp  $\alpha$ -satellite DNA (Figure 5 c). Another specialized centromere strategy is established by some nematodes, like *Caenorhabditis elegans*. So called holocentromeres are spread along the whole chromosome and CENP-A nucleosomes are incorporated into several positions along the chromosome arm (Figure 5 d) (Steiner and Henikoff, 2014).



**Figure 5: Schematic representation of different centromeres**

a) Budding yeast point centromere, this centromere is sequence specific at centromere DNA Elements (CDEI-III)

b+c) regional centromeres can span 4 kb to several megabases (also satellite centromere)

d) some species, like in nematodes, form holocentromeres, centromeres spread over the whole chromosome.

Modified from Steiner & Henikoff, *Current Opinion in Genetics & Development* (2015) (Steiner and Henikoff, 2015)

The only centromere described, that has a specific underlying DNA sequence, is that of *S. cerevisiae*. In all other centromeres any sequence similarity is missing despite sharing common features like repetitive DNA and embedding into pericentric heterochromatin. This led to the hypothesis, that there is no genetic basis of centromere identity but an epigenetic model for centromere identity inheritance.

## 1.4.2 CENP-A determines centromere identity

Three different indications support the hypothesis for an epigenetic model of centromere inheritance. First, the missing sequence similarity between organisms, described in chapter 1.4.1. Second, in chromosomes containing two centromeres, called dicentric chromosomes, one can be inactivated and still maintain the  $\alpha$ -satellite

repetitive sequence (Warburton et al., 1997). Third, the observation that neo-centromeres form *de novo* in chromosome regions lacking any sequence similarity to endogenous centromeric DNA (Amor et al., 2004).

In 1985 it was discovered that CENP-A is essential at centromeres and that it is a major part in centromeric chromatin (Earnshaw and Rothfield, 1985; Palmer and Margolis, 1985). CENP-A is assembled at centromeres in almost all eukaryotic organisms with only a few exceptions. Hence, this centromere specific histone H3 variant is the hallmark for epigenetic specification of centromere identity (Warburton et al., 1997). This hypothesis is confirmed by the presence of CENP-A at neo-centromeres (Marshall et al., 2008) and the circumstance that artificial targeting of CENP-A to ectopic sites is sufficient to generate structures for microtubule recruitment during mitosis for chromosome segregation (Heun et al., 2006; Mendiburo et al., 2011). However it is still an open question how exactly the CENP-A chromatin propagates its identity over many generations.

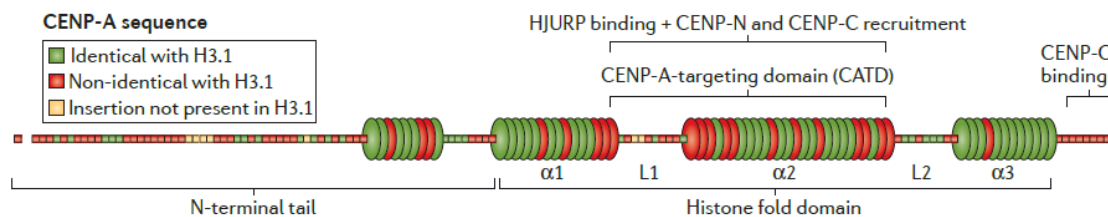
Since CENP-A is the most important factor for centromere inheritance; I will give an overview about its structure and function in section 1.5.

### 1.5 The centromeric histone H3 variant CENP-A

As already mentioned, the CENP-A (or Cse4 in yeast, CID in flies or CenH3 in plants) is a functionally highly conserved protein with distinct functions for kinetochore assembly and centromere identification. In centromeres the CENP-A nucleosomes replace canonical nucleosomes containing H3 and change chromatin structure and features to define the centromere identity (McKinley and Cheeseman, 2016). In addition CENP-A is a binding factor for kinetochore proteins and therefore serves as binding platform for microtubule attachment during mitosis (Schalch and Steiner, 2016). CENP-A consists of two major domains, the histone fold domain with 62 % sequence similarity to H3 and the N-terminal tail with an even higher extent in sequence differences compared to H3 (Sullivan et al., 1994) (Figure 6).

### 1.5.1 CENP-A structural domains

With significant structural and sequence differences compared to the canonical histone H3, CENP-A nucleosomes provide an environment for centromere establishment and centromere identity inheritance. The CENP-A targeting domain (CATD) is included in the histone fold domain and contains the first loop (L1) and the second  $\alpha$ -helix ( $\alpha 2$ ) (Figure 6). This sequence is sufficient for centromere targeting when introduced into H3 histones and therefore the most important domain for defining CENP-A (Black et al., 2007). Targeting and correct localization of CENP-A via the CATD is mediated by its histone chaperone HJURP (holliday junction recognition protein) (Foltz et al., 2009). CATD binds to the N-terminal part of HJURP, whereas the C-terminus of HJURP protects the CENP-A/H4 heterodimer for the formation of a (CENP-A/H4)<sub>2</sub> tetramer by binding the DNA-binding region of the dimer (Hu et al., 2011).



**Figure 6: Primary sequence and secondary structure of CENP-A**

Conservation of CENP-A with H3 is shown in green and non-identical sequence is shown in red. CENP-A has several interaction domains. CATD domain is important for HJURP, the CENP-A chaperone, and CCAN protein recruitment. CENP-C specifically interacts with C-terminal part of CENP-A.

Modified from McKinley & Cheeseman, *Nature Reviews, Molecular Cell Biology* (2016) (McKinley and Cheeseman, 2016)

The CATD is not only important in targeting CENP-A to centromeres, but also in recruitment of the CCAN proteins, CENP-C and CENP-N.

First it was shown that CENP-C recognizes the C-terminal part of CENP-A in frog egg extracts and that this is sufficient to assemble the kinetochore (Guse et al., 2011). Further studies indicated that CATD is important for CENP-C recruitment

(Westhorpe et al., 2015) and the N-terminal tail of CENP-A is involved in enhancing CENP-A/CENP-C interaction (Logsdon et al., 2015).

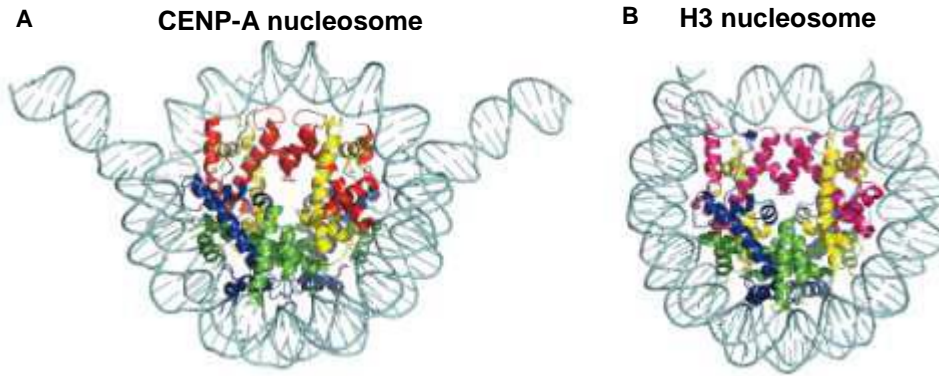
Not only the CENP-A protein, also CENP-A nucleosomes show significant structural differences compared to canonical H3-containing nucleosomes, which will be described in the following.

### 1.5.2 CENP-A nucleosome

Newly synthesized CENP-A/H4 heterodimers are bound and specifically targeted to centromeres by HJURP (Hu et al., 2011). The same conformation of CENP-A/H4 is kept in nucleosome environment (Tachiwana et al., 2011), demonstrating that HJURP is responsible for preventing tetramer formation and DNA association by binding the DNA interactions site.

CENP-A nucleosomes are stabilized by binding of CENP-C into the acidic patch (Kato et al., 2013; Tachiwana et al., 2011). After binding of CENP-C the CENP-A nucleosomes become flattened and more rigid and CENP-A turnover at centromeres is limited (Falk et al., 2015; Falk et al., 2016).

The crystal structure reported in 2011 by Tachiwana et al. (Tachiwana et al., 2011) revealed an octameric structure of CENP-A nucleosomes and summarized four important features of CENP-A nucleosomes. First, the DNA is wrapped left-handed around the histone core. Second, the  $\alpha$ N-helix of CENP-A is shorter and therefore the DNA exit and entry is not fixed equally to H3 nucleosomes (compare Figure 7). 146 bp wrap around H3 nucleosomes (Luger et al., 1997), whereas only ~ 120 bp are tightly wrapped around CENP-A nucleosomes (Sekulic and Black, 2012). Third, the L1 loop, part of the CATD domain is exposed to the surface of the nucleosomes. And fourth, the entire CENP-A nucleosome has a similar overall shape and dimension compared to H3 nucleosomes (Sekulic and Black, 2012).



**Figure 7: Structural model of CENP-A and H3 nucleosome**

A) Less DNA is wrapped around CENP-A nucleosomes since exit and entry site of CENP-A differs from H3 (B).

Adopted from [http://www.yomiuri.co.jp/adv/wol/dy/news/news\\_110721.html](http://www.yomiuri.co.jp/adv/wol/dy/news/news_110721.html)  
23.06.2017

In summary, the (CENP-A/H4)<sub>2</sub> tetramer is more rigid, because the CENP-A/CENP-A interface is rotated compared to H3 nucleosomes, leading to a more compact structure (Sekulic et al., 2010). The CENP-A/CENP-A axis is similar to the H3/H3 axis (Falk et al., 2015) and finally CENP-A nucleosome structure is re-shaped and stabilized upon CENP-C binding (Falk et al., 2015; Falk et al., 2016). This specific CENP-A nucleosome structure influences the centromeric chromatin architecture.

### 1.6 Human centromeric chromatin

As already described in 1.4.1, Figure 5 c, the human centromere is separated into two regions, the centromere core and pericentromeric heterochromatin flanking the centromere core. Disruption of one of these regions leads to defects in chromosome segregation, as shown by functional studies summarized in the review of Schalch and Steiner, 2016 (Schalch and Steiner, 2016). However, it is not completely understood how both regions interact mechanistically and why both regions are so important. The most favored model is that both form a distinct three-dimensional structure of centromeric chromatin that enables microtubule attachment and supports and senses tension for proper chromosome segregation (Schalch and Steiner, 2016).

The centromere core contains the CENP-A nucleosomes and is exposed to surface in condensed mitotic chromosomes. Within the centromere core, the CENP-A nucleosomes are interspersed by H3 containing nucleosomes, shown by quantitative approaches (Bodor et al., 2014). This was also confirmed by super-resolution microscopy of stretched chromatin fibers (Blower et al., 2002). The CENP-A nucleosomes serve as anchoring platform for the kinetochore proteins.

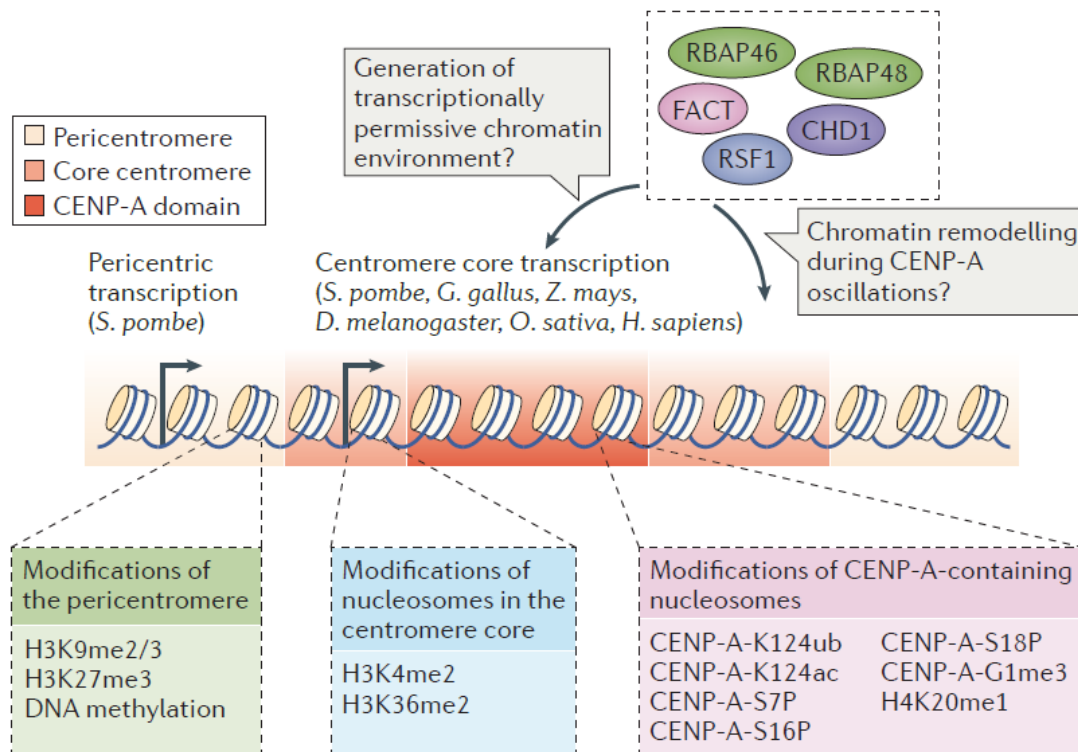
The pericentromeric region is responsible for elasticity and tension resistance generated by cohesin and condensin (Gerlich et al., 2006; Ribeiro et al., 2009). Further function of pericentromeric chromatin is tension sensing and signaling to mitotic checkpoint (Schalch and Steiner, 2016).

The three-dimensional shape of centromere chromatin is dependent on CCAN proteins and centromere specific histone modifications. CCAN proteins, like CENP-T, -W, -S, -X and CENP-B contribute to centromere chromatin architecture as they bind to DNA either sequence specific (CENP-B), to stabilize CENP-A nucleosomes (Fujita et al., 2015) or as nucleosome-like complex (CENP-TWSX) altering DNA conformation (Nishino et al., 2012).

Post-translational histone modifications specify domains of centromere core and pericentromeric chromatin. Within the centromere and pericentromere are histone marks correlating with active chromatin and inactivating histone marks that lead to a more dense packing of DNA respectively.

As represented in Figure 8 (green box), the pericentromeric heterochromatin is highly associated with histone marks leading to inactive chromatin. Hypermethylated H3K9 (H3K9me2/3) and methylated DNA, which both mediate a strong compaction of DNA (Peters et al., 2003), are enriched in pericentromeric chromatin. In addition it is reported that the H4K20 methyltransferase Suv4-20h2 mediates the recruitment of cohesins (Hahn et al., 2013) and loss of H3K9 methylation and HP1 binding increases the separation of major satellites in metaphase (Schalch and Steiner, 2016). H3K27me3 modifications also contribute to inactive chromatin.





**Figure 8: Model of epigenetic modifications on centromeric chromatin**

Pericentromeric chromatin has different post-translational histone modifications than nucleosomes in centromere core and CENP-A nucleosomes themselves. CENP-A nucleosome modifications directly influence centromere function. In several organisms an active transcription of centromere DNA is suggested to be important for centromere function and maintenance.

Adopted from McKinley & Cheeseman, *Nature Reviews, Molecular Cell Biology* (2016) (McKinley and Cheeseman, 2016)

In contrast, the centromere core is more associated with active chromatin marks (Figure 8, blue box). The interspersed H3 nucleosomes are modified on histone 3, lysine 4 with demethylation (H3K4me2) (Blower et al., 2002; Ribeiro et al., 2010; Sullivan and Karpen, 2004). This H3K4me2 modification is functionally important, since HJURP recruitment and CENP-A assembly is dependent on this modifications in a synthetic human kinetochore system (Bergmann et al., 2011). Also the H3K36 dimethylation (H3K36me2) is found at centromeric chromatin in stretches of chromatin fibers and by ChIP analysis (Bergmann et al., 2011).

Furthermore CENP-A nucleosomes themselves are also modified with certain histone marks (Figure 8, red box). For example temporal CENP-A phosphorylations regulate the incorporation of CENP-A histones into centromeric chromatin (McKinley and Cheeseman, 2016). The major modification of CENP-A nucleosomes

is the H4K20 monomethylation (H4K20me1), a histone mark that is essential for kinetochore assembly (Hori et al., 2014).

The CENP-A incorporation into existing centromeres is well known. However, *de novo* centromere establishment and maturation of the centromeric chromatin environment mentioned before are largely unknown. Approaches to study *de novo* centromere establishment are explained in further detail below.

### 1.7 *De novo* centromere formation

Neo-centromeres are centromeres that form at atypical sites on chromosomes, such as chromosome arms or telomeres, and underlying DNA sequences differ from endogenous centromeres (Scott and Sullivan, 2014). Spontaneous *de novo* centromere, or neo-centromere, formation on chromosomes is a relatively rare event, because it results in genomic instability. However, neo-centromeres were described in 1993, associated with abnormal phenotypes, like in cancer cells. Since neo-centromere formation is associated with cancer, several approaches exist to investigate *de novo* centromere formation from scratch.

For example, massive overexpression of CENP-A leads to incorporation of CENP-A throughout the chromosome arm. After an initial pulse of CENP-A expression and its incorporation, CENP-A is removed from most chromosome regions again. In regions where CENP-A stays incorporated, kinetochore proteins are recruited during mitosis and a neo-centromere has formed (Heun et al., 2006). Another method for centromere establishment is artificial targeting of CENP-A to defined DNA loci to induce centromere formation (Mendiburo et al., 2011). By targeting the CENP-A chaperone HJURP to operator sites in human cells, establishment of an ectopic centromere was also achieved (Barnhart et al., 2011).

All these approaches have in common, that the artificial centromere is able to recruit CCAN proteins and the KMN complex. In addition, spindle microtubules are

attached and CENP-A is recruited to these centromeres. However, the major disadvantage of these techniques is that chromosomes are genetically instable after targeting and recruiting CENP-A to ectopic sites, because of chromosome miss-segregation and chromosome break during mitosis (Sekulic and Black, 2012).

Therefore, more suitable approaches for investigation of *de novo* centromere establishment, maturation and propagation are the human artificial chromosomes and artificial targeting of CENP-A to plasmids.

### 1.7.1 Human artificial chromosomes (HACs)

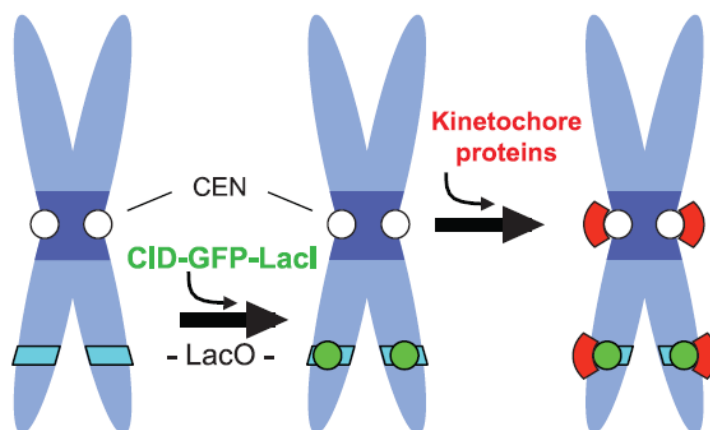
Besides of being interesting vectors for gene delivery and gene therapy (Basu and Willard, 2005), human artificial chromosomes (HACs) are suitable tools to investigate centromere establishment and inheritance. HACs mimic endogenous chromosomes in small scale, since they assemble kinetochores and segregate actively during mitosis (Nakashima et al., 2005). HACs used for centromere studies contain the “alphoid<sup>tetO</sup> array” to enable manipulation of centromeric chromatin in HACs by targeting chromatin modifying enzymes to tetO sites (Nakano et al., 2008). Centromere formation on HACs is dependent on the presence of CENP-B boxes within the alphoid sequences to recruit CENP-B proteins for *de novo* incorporation of CENP-A nucleosomes into the alpha-satellite DNA of the human artificial chromosomes (Masumoto et al., 2004; Ohzeki et al., 2002).

These HACs are used to manipulate chromatin status and to investigate its influence on centromere stability. It was observed that inactive chromatin flanking centromere region is important for HAC formation (Nakashima et al., 2005). By targeting a HAT (histone acetyl-transferase) to tetO within the HACs, Ohzeki et al. found that the balance of H3K9 methylation and acetylation is associated with centromere specification (Ohzeki et al., 2012).

Therefore HACs provide a suitable model system to study and manipulate centromeric chromatin and its influence on kinetochore assembly. However, human artificial chromosomes are huge extrachromosomal DNAs and they have a low transduction efficacy.

### 1.7.2 Centromere formation at ectopic sites on chromosomes

Observations, that overexpression of CENP-A leads to establishment of neo-centromeres and recruitment of kinetochore proteins, gave already hints that CENP-A is sufficient for centromere establishment. However, only a few sites generated ectopic centromeres with this approach, and the direct correlation of CENP-A incorporation and centromere formation was still missing (Heun et al., 2006). To verify the hypothesis that CENP-A generates centromeres, the group of Patrick Heun generated an artificial targeting system. Therefore, a LacO array was introduced into a chromosome of *Drosophila melanogaster* Schneider S2 cells and a CID:GFP:LacI was artificially targeted by pulse induction. CID is the *Drosophila melanogaster* homologue of human CENP-A. Targeting of CID leads to the recruitment of kinetochore proteins, as represented in Figure 9.



**Figure 9: Schematic representation of CENP-A (CID) targeting to LacO sites in chromosomes**

Targeting of CID:GFP:LacI to ectopic sites on chromosomes leads to recruitment of kinetochore proteins to these sites.

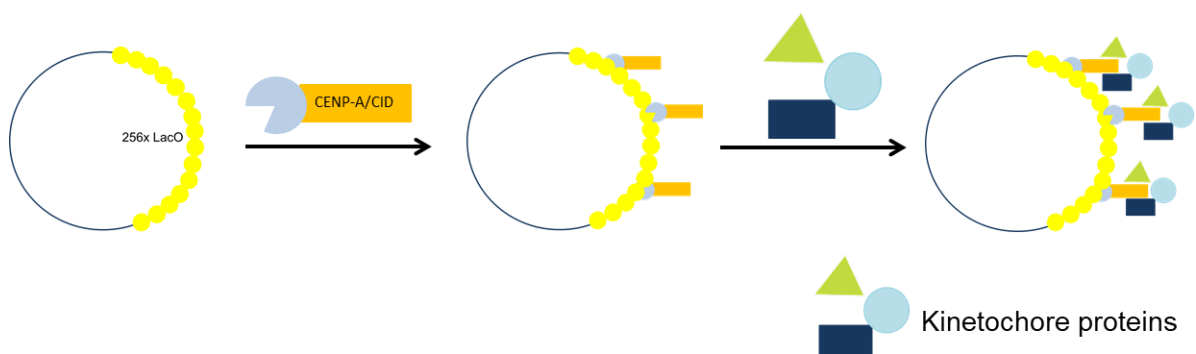
Modified from Mendiburo et al., *Science* (2011) (Mendiburo et al., 2011)

The recruitment of CCAN protein CENP-C and outer kinetochore protein Ndc80 was validated by immune fluorescence of mitotic chromosomes (Mendiburo et al., 2011). As reported in 1941 (McClintock, 1941), dicentric chromosomes cause chromosome break and segregation defects. Since the induction of ectopic centromeres lead to dicentric chromosomes, this method is not suitable to further study centromere maturation and centromere inheritance.

To prevent the genetic instability of *Drosophila melanogaster* Schneider S2 cells, the CID:GFP:LacI fusion protein was targeted to extrachromosomal plasmids harboring an array of 256 LacO sites.

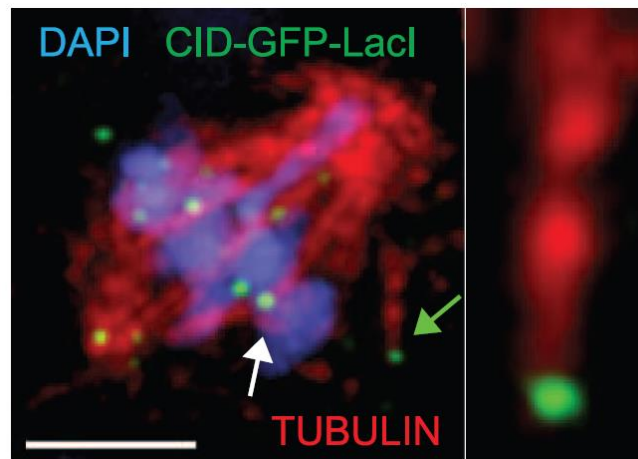
### 1.7.3 *De novo* Centromere formation on LacO plasmids

The current epigenetic model for centromere inheritance is that CENP-A self-directs its loading after each cell division (Allshire and Karpen, 2008). To verify this and to investigate *de novo* centromere heritability, an episomal DNA element containing 256 LacO sequences, a LacO plasmid, was utilized in *Drosophila melanogaster* Schneider S2 cells, as depicted in Figure 10.



**Figure 10: Schematic representation of CENP-A (CID) targeting to LacO plasmids**  
CID:GFP:LacI is targeted to LacO containing plasmids. This leads to the recruitment of kinetochore proteins and to self-propagation of artificial centromeres on plasmids.

Pulse induction of CID:GFP:LacI reveals recruitment of kinetochore proteins (Mendiburo et al., 2011). Expressing a HA tagged CID protein in parallel and performing immune fluorescence on HA and GFP indicated the recruitment of CID to LacO sites independent on LacI-LacO interaction. CID:HA was also represented at the plasmid LacO region. Therefore, by using the plasmids system, not only the *de novo* centromere establishment, but also the inheritance of centromere identity was analyzed.



**Figure 11: Microtubules attach to plasmids**

Immunostaining demonstrates binding of microtubules to plasmids during mitosis.

Modified from Mendiburo et al., *Science* (2011)

(Mendiburo et al., 2011)

Plasmids that are targeted with CID:GFP:LacI establish a centromere identity, which results in the attachment of these plasmids to spindle microtubules during mitosis (Figure 11). This leads to stable maintenance of these plasmids in the cells by active segregation.

In our study in close collaboration with María J. Mendiburo from Patrick Heun's group, we demonstrated that centromeric chromatin identity is inherited after removal of CID:GFP:LacI expression (Mendiburo et al., 2011). In an immune fluorescence experiment episomal plasmids are depicted as CID/CENP-C positive foci during mitosis. Four weeks after transfection, episomal plasmids are still detectable and importantly these are only represented by CID/CENP-C immune fluorescent staining and no longer by CID:GFP:LacI. The conclusion is that initial

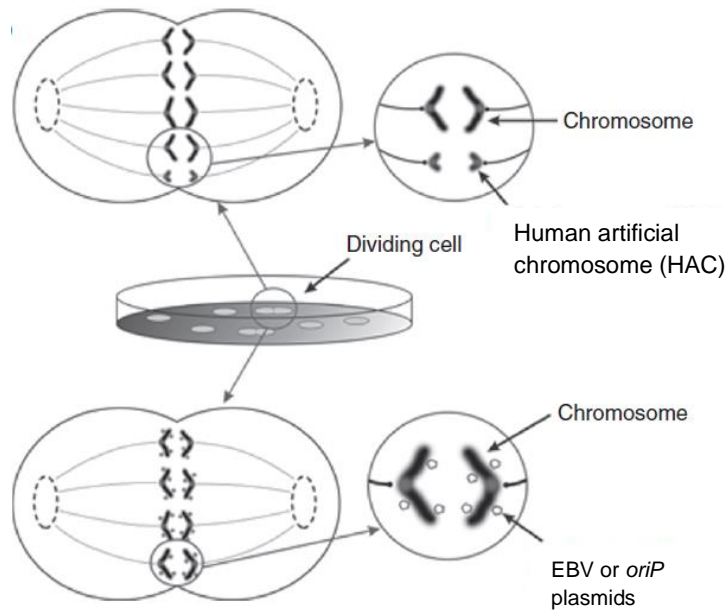
targeting of CID leads to the establishment of centromere identity. This centromere identity is inherited independently of CID targeting over many generations (Mendiburo et al., 2011).

This capacity of self-propagating centromere identity and stable establishment and inheritance of extra-chromosomal DNA make this plasmid system a suitable gene therapy vector. Hence, we transferred the CENP-A targeting to plasmids into a human system. Our group established a tetO-DS reporter plasmid system based on the latent replication origin of the Epstein-Barr virus (EBV) and I investigated early centromere maturation and centromere inheritance.

### 1.8 Development of the pCON<sup>CENP-A</sup> plasmid system

The Epstein-Barr virus (EBV) is a  $\gamma$ -herpesvirus with a double-stranded DNA and a genome size of 172 kbp. It infects resting B-cells, establishes a lifelong persistent infection and is associated with tumor development, like Burkitt's lymphoma or Hodgkin's lymphoma. However, more than 90 % of world's adult population is infected with EBV showing no symptoms of disease (Delecluse and Hammerschmidt, 2000).

Different DNA vectors were generated by cloning the latent replication origin of EBV into plasmids (Pich et al., 2008). First, the latent replication origin of EBV was cloned into bacterial plasmids and these were maintained stable in human cells. *oriP* consists of two cis-acting DNA-elements, the Dyad Symmetry (DS), mediating DNA replication and the Family of Repeats (FR), which is important for plasmid retention during mitosis. Both DNA elements harbor specific binding sites for their trans-activator protein Epstein-Barr virus nuclear antigen 1 (EBNA1). EBNA1 is able to recruit the DNA replication machinery and binds chromosomal DNA for piggybacking EBV genomes during mitosis (Kirchmaier and Sugden, 1995).



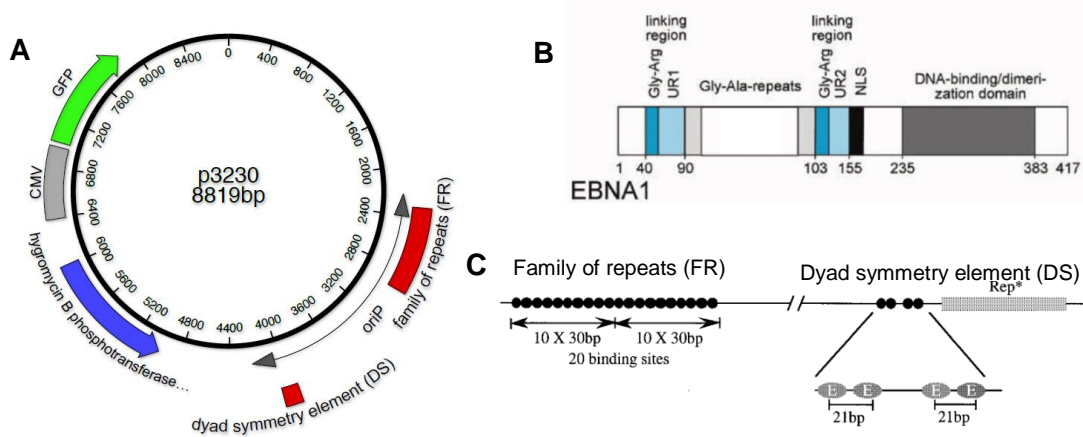
**Figure 12: HAC and *oriP* maintenance during cell division**

During mitosis HACs are segregated actively by attachment to microtubules (top), whereas *oriP* plasmids are associated to host chromosomes and segregated passively by piggybacking.

Modified from Lufino et al, *Molecular Therapy* (2008) (Lufino et al., 2008)

EBV genomes are maintained in infected cells extra-chromosomally and *oriP*-based vectors are also suggested not to integrate and maintained autonomously (Ehrhardt et al., 2008). But other than human artificial chromosomes (HACs), explained in chapter 1.7.1, the *oriP* plasmids are not segregated by an active segregation mechanism during mitosis (compare Figure 12). EBNA1 binds to 20 specific binding sites within the FR and connects plasmids to host chromosome DNA via its DNA binding domain (Yates et al., 2000), a mechanism called piggybacking (Figure 12). The N-terminal part of EBNA1 mediates the binding to AT-rich DNA via the linking regions LR1 and LR2 (Figure 13 B) (Middleton and Sugden, 1992). In addition, the EBNA1 protein has a binding site that specifically detects and binds the DNA sequence present DS in its C-terminal domain (DNA binding domain). EBNA1 binding to DS leads to the recruitment of the origin recognition complex (ORC) and DNA replication (Schepers et al., 2001).





**Figure 13: *oriP* plasmid based on latent replication origin of EBV**

A) *oriP* reporter plasmid: DS element for plasmid replication, FR element for plasmid maintenance during mitosis via piggybacking to host chromosomes, GFP reporter gene and selection marker.

B) EBNA1 protein binds to specific DNA sequence on EBV genome or *oriP* plasmids via N-terminal domain. DNA-binding and dimerization domain of EBNA1 binds to AT-rich sequences on host chromosomes for piggybacking EBV/*oriP* (Adopted from Pich et al., Nucleic acid research (2008) (Pich et al., 2008)

C) bipartite structure of *oriP*: FR 20 binding sites for EBNA1 for plasmid retention and DS with 4 EBNA1 binding sites for plasmid replication via recruiting host replication machinery

Modified from Schepers et al., *EMBO* (2001) (Schepers et al., 2001)

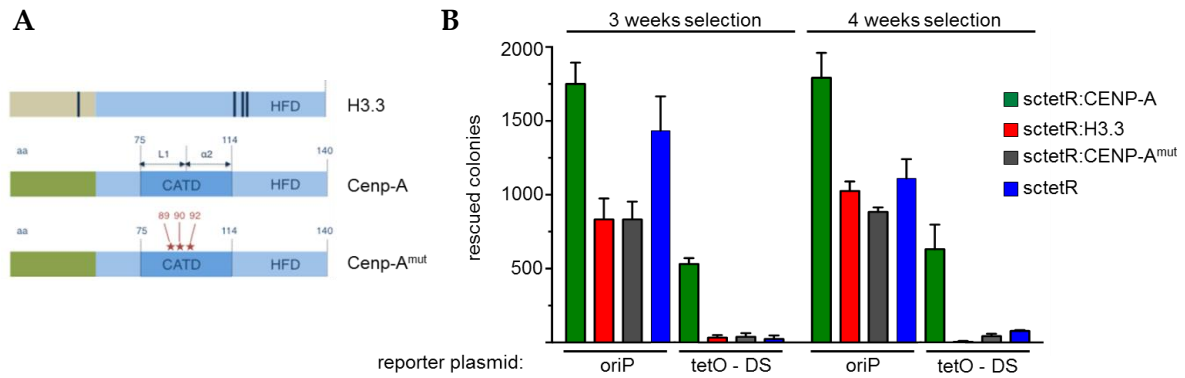
Stable maintenance of *oriP* plasmids is dependent on selection pressure and these plasmids are rapidly lost if selection is removed. Therefore, the *oriP* based plasmid system was further developed to investigate DNA replication and segregation mechanisms independently. For example, our group is examining mechanisms of activating and establishing replication origins in an *oriP*-based plasmid where DS is replaced by a different targeting sequence (Brustel et al., 2017).

Different plasmid segregation mechanisms are examined by the replacement of FR by tetO sites. First, tetO repeats were targeted by the high mobility group protein A (HMGA1). HMGA1a was suggested as a good candidate for plasmid maintenance since it was shown that it interacts with the origin recognition complex (Thomae et al., 2008) and has similar AT-hook domains like the EBNA1 protein. Our group developed a conditional gene vector by fusing HMGA1a to a dimeric single chain (sc) tet transactivator (sctetR). Targeting of sctetR:HMGA1a to tet operator sites lead to

replication and nuclear retention of DS-tetO reporter plasmids. Plasmid loss was induced by an allosteric switch within *sctetR* with doxycycline (Pich et al., 2008). With this system, called pCON<sup>HMGA1a</sup>, our group was successful in demonstrating that HMGA1a has replicative potential and supports plasmid maintenance during cell division similar to EBNA1 by attaching plasmids to host chromosomes (Thomae et al., 2011; Thomae et al., 2008). However, this system is also dependent on selective pressure because plasmids are lost upon removal of selection (Pich et al., 2008).

Therefore, a further development of this system was to target *sctetR*:CENP-A to tetO sites that replace the FR element in the *oriP*, similar to the LacO system, used in *Drosophila melanogaster*.

Stefanie Fülöp, a former group member, investigated the plasmid maintenance in human cells. Stefanie Fülöp demonstrated that just functional CENP-A supports plasmid maintenance.



**Figure 14: tetO-DS plasmid maintenance**

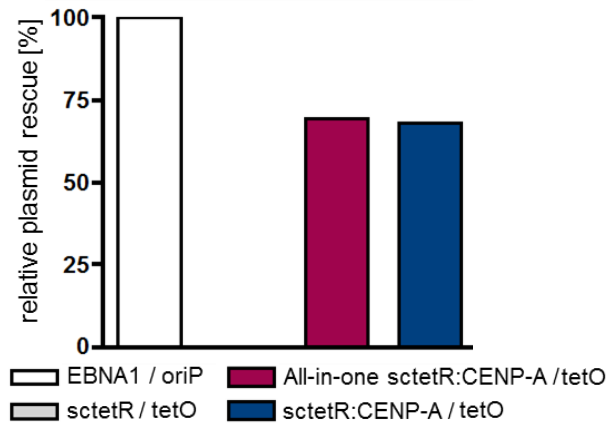
A) Schematic representation of domain differences between H3.3, CENP-A and a mutated version of CENP-A in these experiments are indicated

B) Maintenance of tetO-DS plasmids is dependent on functional CENP-A and does not work with H3.3, mutated CENP-A or *sctetR* targeting. *oriP* is maintained similar in all cell lines, independent of *sctetR* fusion protein

Data: Stefanie Fülöp, *Dissertation* (2013)

Stefanie Fülöp performed a plasmid maintenance experiment in three different cell lines. Cells either expressed *sctetR*:CENP-A, *sctetR*, *sctetR*:H3.3 and *sctetR*:CENP-A<sup>mut</sup>, mutated in the CATD domain of CENP-A. After three or four weeks after

transfection only the targeting of *sctetR*:CENP-A lead to plasmid maintenance of tetO-DS plasmids. *oriP* plasmids were used as control, since the maintenance of *oriP* plasmids is independent of targeting *sctetR* fusion proteins. The *oriP* plasmids were maintained in all cell lines, independent of the expression of different *sctetR* fusion proteins, *sctetR*:CENP-A, *sctetR*:H3.3, *sctetR*:CENP-A<sup>mut</sup> and *sctetR*. In contrast, tetO-DS reporter plasmids were only maintained after targeting of functional CENP-A. The tetO-DS reporter plasmids were not established in the cells when targeted with *sctetR*:H3.3, *sctetR*:CENP-A<sup>mut</sup> or *sctetR* because no centromere identity was established there.

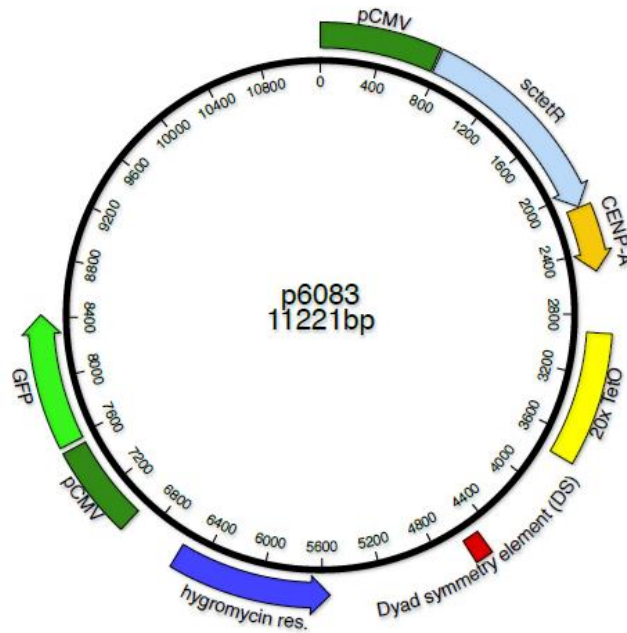


**Figure 15: Plasmid maintenance of different reporter plasmids**

Maintenance of plasmids with targeting of *sctetR*:CENP-A expressed from plasmids (All-in-one) is similar to plasmids where *sctetR*:CENP-A is stably expressed. No plasmid was maintained when *sctetR* alone was targeted.

Data: Stefanie Fülöp, *Dissertation* (2013)

To further develop the tetO-DS *sctetR*:CENP-A plasmid system towards potential gene therapy usage, Stefanie Fülöp analyzed if the expression of the *sctetR*:CENP-A fusion protein in *cis* influences plasmid maintenance. Hence, our group generated an “all-in-one” *sctetR*:CENP-A plasmid and compared its maintenance to *oriP* and the tetO-DS reporter plasmid with *sctetR*:CENP-A expressed in *trans* by stable cell lines. There was no difference in plasmid rescue efficiency between the tetO-DS reporter plasmids with *sctetR*:CENP-A expression in *trans* and the all-in-one plasmid (Figure 15). This “all-in-one” tetO-DS *sctetR*:CENP-A plasmid system we named pCON<sup>CENP-A</sup> (Figure 16).



**Figure 16: All-in-one pCON<sup>CENP-A</sup> reporter plasmid**

sctetR:CENP-A reporter plasmid is based on *oriP*. FR element was replaced by 20 tet operator targeting sites. sctetR:CENP-A is encoded on the plasmid under CMV promoter.

This pCON<sup>CENP-A</sup> plasmid system is an improved vector functional in human cells with potential use for gene therapy approaches. Plasmid maintenance during cell cycle is ensured by active plasmid replication via DS-EBNA1 and the segregation is actively mediated by targeting CENP-A.

During this project our group was investigating the long-term plasmid maintenance and gene expression capacity for using these plasmids for gene therapy approaches.

I used this plasmid system to examine the inheritable centromere established after CENP-A targeting in its functional and epigenetic details.

## 2 AIM OF THE THESIS

In humans, centromere localization is not determined by the underlying DNA sequence. Consequently, its inheritance is specified epigenetically. The major hallmark for centromere identity is the centromere specific histone H3 variant CENP-A. However, the parameters defining centromere identity and its inheritance, such as histone modifications or CENP-A itself, are unclear.

The pCON<sup>CENP-A</sup> plasmid system is a suitable tool for analyzing centromere-dependent mechanisms, as site-specific targeting of CENP-A to the plasmid leads to the formation of a functional neo-centromere. It is stably maintained in human cells because CENP-A targeting leads to active plasmid segregation during mitosis, which makes it a potential tool for gene therapy.

Using the pCON<sup>CENP-A</sup> plasmid system, I aim to understand the molecular mechanisms that lead to the establishment of centromere identity and its inheritance. I am addressing three distinct aspects:

- I) Do pCON<sup>CENP-A</sup> plasmids segregate by a CENP-A-dependent active plasmid segregation mechanism?
- II) What is the minimal plasmid establishment time required for long-term maintenance and CENP-A self-propagation?
- III) Are histone modifications present on a mature plasmid centromere?

I) Established pCON<sup>CENP-A</sup> plasmids segregate actively during cell division. To visualize and follow this process, live cell imaging of cells containing pCON<sup>CENP-A</sup> will be performed. Furthermore, I will analyze artificial centromeres of pCON<sup>CENP-A</sup> for their capacity to recruit kinetochore components by immune fluorescence.

II) My laboratory already showed that CENP-A targeting leads to stable plasmid maintenance after an initial plasmid establishment of three weeks. This plasmid maintenance becomes independent of CENP-A targeting. I aim to examine the minimal timeframe of centromere maturation in which centromeres begin to self-

propagate. Additionally, immune fluorescence experiments will show incorporation of targeting-independent CENP-A.

III) To investigate the histone modifications present on matured artificial centromeres, I will develop a Cas9-dependent targeting system to purify plasmids with matured centromeres from human cells. By Western Blot and mass spectrometry, I aim to reveal histone modifications that are established at centromeres during its maturation.

In conclusion, I aim to understand the molecular mechanisms leading to the establishment of an inheritable centromere and the functional relevance of these artificial centromeres on plasmids for potential gene therapy vectors.

### 3 MATERIALS & METHODS

#### 3.1 Materials

In the following the materials, like devices, chemicals, enzymes and buffers, which were used during this work are listed. Buffers were mixed from stock solutions, prepared from the listed chemicals.

##### 3.1.1 Devices and consumables

In Table 1 devices and consumables are listed in alphabetical order with reference to the distributor. Cell culture devices, like cell culture dishes, cryotubes or 6-well plates were used from Nunc (Nunc GmbH, Germany).

Table 1 Devices and consumables used in this work

Devices	Distributor
Axiovert 10 fluorescence microscope	Carl Zeiss AG, Germany
CEA Blue sensitive X-ray film	Agfa HealthCare GmbH, Germany
Cell counting chamber Neubauer improved	Brand GmbH & Co KG, Germany
Cell Strainer (70 µm)	BD Falcon™, USA
Centrifuge Tubes (15 ml, 50 ml)	Greiner Bio-One International AG, Austria
Centrifuge Rotina 38R	Hettich GmbH & Co. oHG, Germany
Coverslip	Carl Roth GmbH & Co KG, Germany
Developer machine Optimax®	Typon Röntgen-Film GmbH, Germany
Electroporation cuvettes 1mm	Peqlab Biotechnologie GmbH, Germany
Electroporation device	Bio-Rad Laboratories GmbH, Germany
Gene pulser II	
FACS tubes	BD Falcon™, USA
5 ml polypropylene tubes	
Film cassettes	Fujifilm Holdings K.K., Japan
Flip-cap tubes	Sarstedt AG & Co, Germany
Flow Cytometry Analyzer FACSCalibur	BD Biosciences™, USA
Freezer (-20 °C)	Liebherr-International Deutschland GmbH, Germany
Freezer (-80 °C)	Azbil Telstar, S.L.U., Spain
Freezing box	Nunc GmbH, Germany
Nalgen Nunc Cryo 1 °C Freezing Container	

### 3 MATERIALS & METHODS

French Pressure Cell Press French Press FA-078A With 40K standard cell (diameter 3")	Thermo Electron GmbH, Germany
Fridge (4 °C)	Liebherr-International Deutschland GmbH, Germany
Gel electrophoresis system	peqlab GmbH, Germany
Gene Amp PCR System 2400	PerkinElmer Inc., USA
High-speed centrifuge Avanti J-26XP	Beckman Coulter GmbH, Germany
Hotplate/ magnetic stirrer RH basic	IKA Labortechnik GmbH & Co KG, Germany
Incubator 37 °C, 5 %CO <sub>2</sub>	UniEquip Laborgerätebau- und Vertriebs GmbH, Germany
Mammalian cells Incubator 37 °C	Heraeus GmbH, Germany
Bacteria Laminar flow	Heraeus GmbH, Germany
LaminAir Hb 2448	
Leica confocal microscope Leica TCS SP2	Leica Microsystems GmbH, Germany
Leica inverted motorized live cell fluorescence microscope Leica DMI8 (BMC Munich)	Leica Microsystems GmbH, Germany
Light-Cycler®	Roche Diagnostics GmbH, Germany
Live cell imaging chamber µ-Slide 8 Well; ibiTreat	Ibidi GmbH, Germany
Microscopy slide SuperFrost®	Carl Roth GmbH & Co KG, Germany
Millipore Water purification system Milli-RO 60 plus	Merck Millipore, Merck KGaA, Germany
NanoDrop® ND-1000 spectrometer	Thermo Fisher Scientific Inc., USA
Nitrocellulose ECL blotting membrane, Amersham	GE Healthcare GmbH, Germany
Orbital shaker Innova 4400	New Brunswick Scientific GmbH, Germany
Parafilm® M	Brand GmbH & Co KG, Germany
Photometer	Eppendorf AG, Germany
Pipet-boy	IBS Integra Biosciences GmbH, Germany
Pipet tips 10 µl, 200 µl, 1000 µl	Gilson Inc., USA
Qubit-Fluorometer	Invitrogen GmbH, Germany
SemiDry Blotting System	Hoefer Scientific Instruments, USA
Sterile filter 0.45 µm, syringe filter	Sartorius AG, Germany
Sonifier Covaris S220	Covaris Inc., Great Britain
Sonifier tubes	Covaris Inc., Great Britain
AFA Fiber & Cap tubes (12x12 mm)	
Syringe (1 ml to 50 ml)	Norm-Ject, Henke-Sass, Wolf GmbH, Germany
Table centrifuge 5415R	Eppendorf AG, Germany
Table centrifuge	A. Hartenstein GmbH, Germany
Thermomixer comfort	Eppendorf AG, Germany
Vortex Genie 2	Scientific Industries Inc., USA



#### 3.1.2 Software

Software-programs and their developers are listed in Table 2.

Table 2 Software used during this work

Software	Developer
EndNote X7	Clarivate Analytics, USA
FACSDIVA™ V6.1.1	BD Biosciences™, USA
Fiji	ImageJ, USA
FileMaker Pro 15	FileMaker Inc., USA
FlowJo 10.0.8r1	FlowJo LLC, USA
Leica Application Suite	Leica Microsystems GmbH, Germany
LightCycler® 480 Software 1.5.162 SP2	Roche Diagnostics GmbH, Germany
MacVector 13.5.5	MacVector Inc., USA
MS Office 2010	Microsoft Corporation, USA
Prism 6.0c	GraphPad Software Inc., USA

#### 3.1.3 Enzymes and antibodies

In Table 3 enzymes are specified. Restriction enzymes used for cloning and DNA fragmentation for the STREP pull down (chapter 3.2.9) were purchased from New England Biolabs Inc., Great Britain.

Table 3 Enzymes used in this work

Enzyme	Distributor
Alkaline Phosphatase, Calf Intestine (CIP)	New England Biolabs Inc., Great Britain
MNase	Roche Diagnostics GmbH, Germany
Proteinase K	Roche Diagnostics GmbH, Germany
Pwo Polymerase	peqlab GmbH, Germany
RNase A, DNase free	Roche Diagnostics GmbH, Germany
T4 DNA ligase	Affymetrix Inc., USA

Antibodies, their specification and their application and dilution during this work are given in Table 4 (ChIP, Co-IP and Western Blot) and Table 5 (Immune fluorescence).

### 3 MATERIALS & METHODS

Table 4 Antibodies used in ChIP, Co-IP and Western Blot

Antibody	Origin	Application	Dilution	Distributor
<b><math>\alpha</math>-tetR (31B3)</b>	Mouse	ChIP; Co-IP WB	See text 1:20	Helmholtz Center Munich
<b><math>\alpha</math>-HA (12CA5)</b>	Mouse	ChIP WB	See text 1:20	Helmholtz Center Munich
<b><math>\alpha</math>-H3 (ab1791)</b>	Rabbit	WB	1:5000	Abcam plc, Great Britain
<b><math>\alpha</math>-H2B (ab1790)</b>	Rabbit	WB	1:1000	Abcam plc, Great Britain
<b><math>\alpha</math>-strep (12B8)</b>	Rat	WB	1:20	Helmholtz Center Munich
<b><math>\alpha</math>-H3K4me1 (ab8895)</b>	Rabbit	WB	1:500	Abcam plc, Great Britain
<b><math>\alpha</math>-H3K4me2 (ab7766)</b>	Rabbit	WB	1:1000	Abcam plc, Great Britain
<b><math>\alpha</math>-H3K4me3 (ab8580)</b>	Rabbit	WB	1:1000	Abcam plc, Great Britain
<b><math>\alpha</math>-H3K9me2 (ab1220)</b>	Mouse	WB	1:200	Abcam plc, Great Britain
<b><math>\alpha</math>-H3K9me3 (ab8898)</b>	Rabbit	WB	1:1000	Abcam plc, Great Britain
<b><math>\alpha</math>-H3K27me1 (07-448)</b>	Rabbit	WB	1:2000	Merck Millipore, Merck KGaA, Germany
<b><math>\alpha</math>-H3K27me2 (07-425)</b>	Rabbit	WB	1:1000	Merck Millipore, Merck KGaA, Germany
<b><math>\alpha</math>-H3K27me3 (#9733)</b>	Rabbit	WB	1:500	Cell signaling Technology Inc., USA
<b><math>\alpha</math>-H4K20me1 (C15410034)</b>	Rabbit	WB	1:1000	Diagenode Inc., USA
<b><math>\alpha</math>-H4K20me2 (9759S)</b>	Rabbit	WB	1:1000	Cell signaling Technology Inc., USA
<b><math>\alpha</math>-H4K20me3 (C15410057)</b>	Rabbit	WB	1:1000	Diagenode Inc., USA
<b><math>\alpha</math>-mouse HRP</b>	Goat	WB	1:10000	Cell signaling Technology Inc., USA
<b><math>\alpha</math>-rat HRP</b>	Goat	WB	1:10000	Jackson ImmunoResearch Laboratories Inc., USA
<b><math>\alpha</math>-rabbit HRP</b>	Goat	WB	1:10000	Jackson ImmunoResearch Laboratories Inc., USA

### 3 MATERIALS & METHODS

Table 5 Antibodies used for immune fluorescence in this work

Antibody	Origin	Application	Dilution	Distributor
$\alpha$ -EBNA1 (1H4)	Rat	IF	1:20	Helmholtz Center Munich
$\alpha$ -CENP-C (ab50974)	Mouse	IF	1:100	Abcam plc, Great Britain
$\alpha$ -HEC1 (Ndc80) (ab3613)	Mouse	IF	1:100	Abcam plc, Great Britain
$\alpha$ -rat Cy3	Goat	IF	1:100	Jackson ImmunoResearch Laboratories Inc., USA
$\alpha$ -rat A647	Goat	IF	1:100	Jackson ImmunoResearch Laboratories Inc., USA
$\alpha$ -mouse Cy5	Goat	IF	1:100	Jackson ImmunoResearch Laboratories Inc., USA

#### 3.1.4 Chemicals

In Table 6 chemicals and substances used during this work and their corresponding distributor are listed in alphabetical order.

Table 6 Substances used in this work

Substance	Distributor
Agarose, UltraPure	Invitrogen GmbH, Germany
Ammoniumperoxodisulfate (APS)	Carl Roth GmbH & Co KG, Germany
Ampicillin sodium salt	Carl Roth GmbH & Co KG, Germany
ATX Ponceau S red staining solution	Fluka® Analytical, Germany
Bio-Rad Protein Assay	Bio-Rad Laboratories GmbH, Germany
Bovine Serum Albumin, BSA	Sigma-Aldrich Chemie GmbH, Germany
Calcium chloride (CaCl <sub>2</sub> )	Merck-Eurolab GmbH, Germany
Chlorophorm	Merck-Eurolab GmbH, Germany
Complete protease inhibitor EDTA free	Roche Diagnostics GmbH, Germany
Hydrogen peroxide solution	Sigma-Aldrich Chemie GmbH, Germany
Dimethylpimelimidate (DMP)	Sigma-Aldrich Chemie GmbH, Germany
Dimethylsulphoxide (DMSO)	Carl Roth GmbH & Co KG, Germany
Dithiothreitol (DTT)	Sigma-Aldrich Chemie GmbH, Germany
Deoxycholic acid (DOC)	Sigma-Aldrich Chemie GmbH, Germany
dNTPs	Roche Diagnostics GmbH, Germany
Doxycycline	Sigma-Aldrich Chemie GmbH, Germany
Dulbecco's Eagle Modified Medium	Gibco, Thermo Fisher Scientific Inc., USA
Ethylenediaminetetraacetic acid (EDTA)	Carl Roth GmbH & Co KG, Germany
Ethylenglycol-bis( $\beta$ -aminoethylether)- tetraacetic acid (EGTA)	Carl Roth GmbH & Co KG, Germany

### 3 MATERIALS & METHODS

<b>Ethanol</b>	Panreac AppliChem GmbH, Germany
<b>Ethanolamine</b>	Carl Roth GmbH & Co KG, Germany
<b>Ethidiumbromide</b>	Carl Roth GmbH & Co KG, Germany
<b>Fetal bovine serum (FBS)</b>	Lot BS225160.5, Bio&SELL GmbH, Germany
<b>Formaldehyde 16 %, methanol-free</b>	Thermo Fisher Scientific Inc., USA
<b>G418/ Geneticin</b>	Carl Roth GmbH & Co KG, Germany
<b>GeneRuler™ 1 kb DNA Ladder</b>	Thermo Fisher Scientific Inc., USA
<b>GeneRuler™ DNA Ladder Mix</b>	Thermo Fisher Scientific Inc., USA
<b>Glycerol</b>	AppliChem GmbH, Germany
<b>Glycin</b>	Panreac AppliChem GmbH, Germany
<b>HEPES 100x solution</b>	Gibco, Thermo Fisher Scientific Inc., USA
<b>Hygromycin</b>	PAA-Laboratories, Austria
<b>Immersion oil</b>	Leica Microsystems GmbH, Germany
<b>TypeF Immersion liquid n=1.5180</b>	
<b>Isoamyl alcohol</b>	Merck-Eurolab GmbH; Germany
<b>Isopropanol (2-Propanol)</b>	Merck-Eurolab GmbH; Germany
<b>Kanamycin</b>	Carl Roth GmbH & Co KG, Germany
<b>Light-Cycler®Fast-Start-DNA-Master-SYBR-Green-I</b>	Roche Diagnostics GmbH, Germany
<b>Lipofectamine 2000</b>	Invitrogen GmbH, Germany
<b>L-<math>\alpha</math>-lyso-lecitine</b>	Calbiochem, EMD Biosciences Inc., USA
<b>Manganese chloride (MnCl<sub>2</sub>)</b>	Fluka® Analytical, Germany
<b>2-Mercaptoethanol</b>	Carl Roth GmbH & Co KG, Germany
<b>Methanol</b>	Merck-Eurolab GmbH; Germany
<b>Milk powder</b>	Merck KGaA, Germany
<b>NP-40 (Igepal CA-630)</b>	Sigma-Aldrich Chemie GmbH, Germany
<b>NEB Buffer 1.1; 2.1; 3.1 and Cut Smart</b>	New England Biolabs Inc., Germany
<b>Opti-MEM®</b>	Gibco, Thermo Fisher Scientific Inc., USA
<b>Orange G</b>	Sigma-Aldrich Chemie GmbH, Germany
<b>PageRuler Plus Prestained Protein Ladder</b>	Thermo Fisher Scientific Inc., USA
<b>Penicillin/Streptomycin 100x</b>	Gibco, Thermo Fisher Scientific Inc., USA
<b>Phenol</b>	Carl Roth GmbH & Co KG, Germany
<b>piperazine-N,N'-bis(2-ethanesulfonic acid) (PIPES)</b>	Sigma-Aldrich Chemie GmbH, Germany
<b>Polyacrylamide</b>	Carl Roth GmbH & Co KG, Germany
<b>Polyethylenimin (PEI)</b>	Polyscience Europe GmbH, Germany
<b>Poly-L-Lysine</b>	Sigma-Aldrich Chemie GmbH, Germany
<b>Potassium chloride</b>	ICN Biomedicals Inc., Germany
<b>protein G Sepharose 4 Fast Flow</b>	GE Healthcare, Germany
<b>Puromycin</b>	AppliChem GmbH, Germany
<b>Salmon Sperm</b>	Invitrogen GmbH, Germany
<b>Select Agar</b>	Invitrogen GmbH, Germany
<b>SiR-tubulin</b>	Spirochrome AG, Switzerland
<b>Sodium acetate</b>	Merck-Eurolab GmbH, Germany
<b>Sodium azide</b>	Sigma-Aldrich Chemie GmbH, Germany
<b>Sodium borate</b>	Sigma-Aldrich Chemie GmbH, Germany

### 3 MATERIALS & METHODS

<b>Sodium dodecylsulfate (SDS)</b>	Serva Electrophoresis GmbH; Germany
<b>Sodium chloride</b>	Merck-Eurolab GmbH, Germany
<b>Sodium hydroxide</b>	Carl Roth GmbH & Co KG, Germany
<b>Strep-Tactin® Sepharose®</b>	Iba GmbH, Germany
<b>T4 DNA Ligase Buffer (10x)</b>	Affymetrix Inc., USA
<b>Thymidine</b>	Sigma-Aldrich Chemie GmbH, Germany
<b>TEMED</b>	Carl Roth GmbH & Co KG, Germany
<b>Tris-Base</b>	AppliChem GmbH, Germany
<b>Tris-EDTA pH 8.0</b>	Carl Roth GmbH & Co KG, Germany
<b>Triton-X100</b>	Sigma-Aldrich Chemie GmbH, Germany
<b>Trypsin-EDTA</b>	Gibco, Thermo Fisher Scientific Inc., USA
<b>Tween-20</b>	AppliChem GmbH, Germany
<b>VectaShield™ Mounting Medium</b>	Vector Laboratories Inc., USA
<b>Xfect™ transfection reagent</b>	Clontech, Takara Bio USA Inc., USA

#### 3.1.5 Kits

Kits used during this project are listed in Table 7.

Table 7 Kits used in this work

<b>Kit</b>	<b>Distributor</b>
<b>JetStar 2.0 plasmid purification kit</b>	Genomed GmbH, Germany
<b>NucleoSpin Gel and PCR Clean-up kit</b>	Macherey-Nagel GmbH & Co KG, Germany
<b>CloneJET PCR Cloning Kit</b>	MBI Fermentas, Thermo Fisher Scientific Inc., Germany

#### 3.1.6 Buffer

Buffer compositions of protocols performed during this work are indicated from Table 8 to Table 13.

Table 8 Buffer for generation of chemical competent DH5α

<b>Buffer name</b>	<b>Buffer composition</b>
<b>Inoue Wash Buffer</b>	55 mM MnCl <sub>2</sub> , 15 mM CaCl <sub>2</sub> , 250 mM KCl, 10 mM PIPES
<b>Inoue Buffer</b>	Inoue Wash Buffer, DMSO

### 3 MATERIALS & METHODS

Table 9 Buffer for plasmid rescue assay

Buffer name	Buffer composition
TEN Buffer	10 mM Tris-HCl pH 7.5, 1 mM EDTA, 150 mM NaCl
2xHIRT Buffer	20 mM Tris-HCl pH 7.5, 20 mM EDTA, 1.2 % SDS

Table 10 Buffer for cell lysis and Western Blot

Buffer name	Buffer composition
RIPA extract Buffer	50 mM Tris-HCl pH 7.9, 150 mM NaCl, 1 % NP-40, 0.2 % SDS
5xLaemmli Buffer	250 mM Tris pH 6.8, 10 % SDS, 500 mM DTT, 25 % Glycerol, 0.5 % Bromphenolblue
8 % polyacrylamide separation gel	8 % polyacrylamide, 3.4 mM SDS, 375 mM Tris pH 8.8
13 % polyacrylamide separation gel	13 % polyacrylamide, 3.4 mM SDS, 37 mM Tris pH 8.8
Stacking gel	4 % polyacrylamide, 3.4 mM SDS, 125 mM Tris pH 6.8
1xRunning Buffer	192 mM Glycine, 3.4 mM SDS, 24 mM Tris pH 7.4
Blotting Buffer	1xRunning Buffer, 20 % MeOH
ECL developing solution	1 ml <b>solution A</b> (100 mM Tris pH 8.8, 200 mM p-cumaric acid, 1.25 mM Luminol), 3 µl <b>solution B</b> (3 % (v/v) H <sub>2</sub> O <sub>2</sub> ) :

Table 11 Buffer for covalent coupling of antibodies to protein A/G sepharose beads

Buffer name	Buffer composition
Sodium borate Buffer	0.2 M sodium borate pH 9.0
DMP Buffer	Sodium borate Buffer, 20 mM DMP
Ethanolamine Buffer	0.2 M ethanolamine pH 8.0
Sodium azide Buffer	PBS, 0.01 % sodium azide

Table 12 Buffer for Co-precipitation

Buffer name	Buffer composition
Permeabilizing buffer	50 mM Tris pH 7.9, 10 mM NaCl, 2 mM CaCl <sub>2</sub> , 150 mM sucrose
Lysis buffer	10 mM Tris pH 6.8, 5 mM EDTA, 600 mM NaCl, 1x complete protease inhibitor
Dilution buffer	10 mM Tris pH 6.8, 5 mM EDTA, 1x complete protease inhibitor
LiCl	250 mM LiCl, 0.1 % SDS, 0.5 % DOC, 1 % NP-40, 50 mM Tris pH 8.0, 1 mM EDTA
RIPA-300	300 mM NaCl, 0.1 % SDS, 0.5 % DOC, 1 % NP-40, 50 mM Tris pH 8.0, 1 mM EDTA
RIPA-150	150 mM NaCl, 0.1 % SDS, 0.5 % DOC, 1 % NP-40, 50 mM Tris pH 8.0, 1 mM EDTA
1xTE	Tris-EDTA pH 8.0

### 3 MATERIALS & METHODS

Table 13 Buffer for ChIP and STREP pull down

Buffer name	Buffer composition
2% Formaldehyde	PBS, 2 % Formaldehyde
LB3(+) Buffer	25 mM Hepes pH 7.5, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.5 % Sarcosyl, 0.1 % DOC, 0.5 % Triton-X-100, 1x complete protease inhibitor
Blocking Solution	0.5 mg/ml BSA, 30 µg/ml salmon sperm, 1x complete protease inhibitor, 0.1 % Triton-X-100 in LB3(-) buffer (without detergents)
LiCl	250 mM LiCl, 0.1 % SDS, 0.5 % DOC, 1 % NP-40, 50 mM Tris pH 8.0, 1 mM EDTA
RIPA-300	300 mM NaCl, 0.1 % SDS, 0.5 % DOC, 1 % NP-40, 50 mM Tris pH 8.0, 1 mM EDTA
RIPA-150	150 mM NaCl, 0.1 % SDS, 0.5 % DOC, 1 % NP-40, 50 mM Tris pH 8.0, 1 mM EDTA
1xTE	Tris-EDTA pH 8.0

#### 3.1.7 Plasmids

Plasmids used for ChIP, IF and plasmid maintenance experiments are given in Table 14.

Table 14 Reporter plasmids used for ChIP, IF and plasmid maintenance experiments

Plasmid (AGV identification)	Plasmid description
3279	CMV-EBNA1 <sup>AGA</sup>
3230	FR-DS + GFP; <i>oriP</i>
3293	40xtetO-DS + GFP
3448	20xtetO-DS + GFP + sctetR:HMGA1a
6083	20xtetO-DS + GFP + CMV-sctetR:CENP-A; pCON <sup>CENP-A</sup>
5600	40xtetO-DS + sctetR:GFP:CENP-A; pCON <sup>CENP-A</sup>
5602	40xtetO-DS + sctetR:GFP
4292	20xtetO-DS + GFP + miniEcad-sctetR:CENP-A; pCON <sup>CENP-A</sup>
6437	20xtetO-DS-40xCas9 targ. + GFP + sctetR:CENP-A; pCON <sup>CENP-A</sup> + Cas9 targeting
6359	40xtetO-DS-10xCas9 targ. + sctetR:GFP:CENP-A; pCON <sup>CENP-A</sup> + Cas9 targeting

Plasmids used for cloning the CRIPR/Cas9 targeting system are listed in Table 15.

Table 15 Plasmids used for CRISPR/Cas9 cloning

Plasmid (AGV identification)	Plasmid description
6350	sgRNA A66407 expressing plasmid
6351	10xCas9 targeting sequence (GeneScript)
6372	20xCas9 targeting sequence
6378	40xCas9 targeting sequence
6339	CMV-Cas9:3xmCherry (addgene #64108)
4202	MiniEcad promoter
6367	MiniEcad-Cas9:3xmCherry
6395	CMV-Cas9:mCherry:TAP

## 3.2 Methods

In this chapter detailed information about experimental procedures during this work is given.

### 3.2.1 Molecular biological methods

This chapter deals with cloning techniques, PCR amplification and plasmid purification from bacteria.

#### 3.2.1.1 *Cloning strategies*

##### Multimerization of Cas9 targeting sites

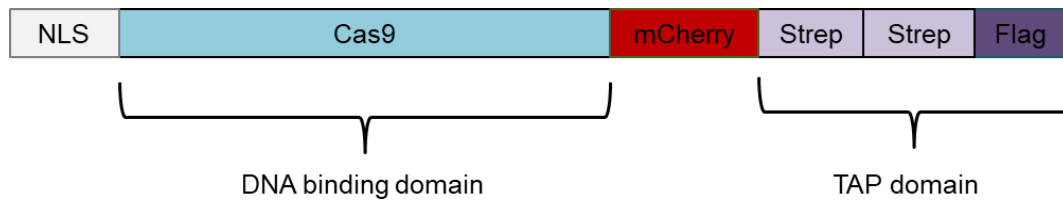
For multimerization of Cas9 targeting sites the decamer plasmid (6351, see Table 15) was digested with BamHI and BglII. The same plasmid was digested with BamHI for linearization (= vector) and with BamHI + BglII for isolating the 10x Cas9 targeting sequence (= insert). Ligation of the decamer sequence and the linearized plasmid with the decamer sequence led to the 20x Cas9 targeting sequence plasmid. 30x and



40x Cas9 targeting sequence was generated by BamHI and BglII digest as well. After generating different numbers of repeats, those were cloned into reporter plasmids by restriction sites SpeI, KpnI or BssHII next to tetO repeats.

#### Cloning of Cas9:mCherry:TAP

The Cas9:3xmCherry plasmid was ordered from addgene (addgene #64108). It was used for multicolor imaging of chromosomal loci by Ma et al. (Ma et al., 2015). The TAP tag, consisting of a tandem STREP and a Flag tag, was cloned onto the Cas9 protein of this Cas9:3xmCherry plasmid. Final fusion protein was generated as depicted in Figure 17.



**Figure 17 schematic representation of Cas9:mCherry:TAP fusion protein**

The Cas9 fusion protein consists of a nuclear localization signal (NLS) to be transported into the nucleus for binding DNA. Cas9 domain binds, together with sgRNA, the target DNA. mCherry is for visualization of the protein within the nucleus. TAP tag consists of tandem STREP and Flag.

For cloning STREP-STREP-Flag onto Cas9:mCherry a PCR was performed to amplify STREP-STREP-Flag with specific restriction sites (see Table 16), NotI and XhoI.

**Table 16 Primer for PCR for cloning Cas9:mCherry:TAP**

Primer name	Sequence
Cas9_for (783)	gagagcggccgcagctggagccaccctcagttc
Cas9_rev (785)	tctcctcgagatttcgaaattcattatcatcatcatctttataatcctctcc

After PCR, pJET1.2 ligation and bacterial amplification of pJET1.2 plasmid with PCR product, the vector plasmid, Cas9:3xmCherry plasmid (6339, Table 15), was digested with NotI + XhoI and the insert plasmid, STREP-STREP-Flag in pJET1.2 was also digested with NotI + XhoI. Both were ligated and transformed into bacteria.

#### 3.2.1.2 Polymerase Chain Reaction (PCR)

Polymerase chain reaction (PCR) is used for specific amplification of DNA fragments (Mullis et al., 1986). Specific primer pairs, designed to amplify the TAP tag from a template plasmid were used. The reaction with *pwo* polymerase was performed in 50 µl total volume and with 50 pg template DNA according to manufacturer's instructions. Temperature profile used for the reaction is shown in Table 17.

Table 17 PCR program for cloning Cas9:mCherry:TAP

	Temperature	Duration	Cycles
<b>Pre-incubation</b>	94 °C	5 min	1
<b>Amplification</b>	94 °C	30 sec	25
	59 °C	30 sec	
	68 °C	45 sec	
<b>Amplification</b>	68 °C	5 min	1
<b>Cooling</b>	4 °C	∞	1

The annealing temperature was calculated according to the respective primer melting temperature. PCR products were gel purified by the NucleoSpin Gel and PCR clean up kit and cloned into the pJET1.2 vector according to manufacturer's instructions.

#### 3.2.1.3 *Generation of chemical competent DH5α bacteria*

Generation of chemical competent DH5α was done as described in Cold Spring Harbor protocol of Sambrook et al. (Sambrook and Russell, 2006)

In brief, a stock of previously prepared competent DH5α was used to inoculate a new 25 ml starting culture. Bacteria were incubated on an orbital shaker for six to eight hours at 37 °C and 200 rpm. Overnight incubation of a 250 ml culture was performed by inoculation 1 ml of the starting culture in 250 ml pre-cooled LB medium and keeping them on an orbital shaker at 18 °C and 200 rpm. The next day the culture was diluted to an OD600 of 0.05 in a total volume of 400 ml. After another overnight incubation at 18 °C and 200 rpm, OD600 measurement was performed continuously until an OD600 of 0.55 was reached. Bacteria were kept on ice-water for 10 min and then harvested by centrifugation at 2500 g for 10 minutes at 4 °C. Medium was discarded and cells were resuspended in 80 ml ice-cold Inoue wash buffer. Cells were centrifuged as described before and resuspended in 20 ml of ice-cold Inoue buffer supplemented with 1.5 ml DMSO. After incubation for 10 min at 4 °C aliquots of 500 µl bacteria suspension were prepared. Bacteria were stored at -80 °C until use after snap-freezing the aliquots in liquid nitrogen.

#### 3.2.1.4 *Transformation*

For transformation chemical competent *Escherichia coli* strain DH5α were thawed on ice and 100 µl of bacteria were mixed with plasmid DNA. The mixture was incubated 30 minutes on ice, transferred to 42 °C for 90 seconds and stored on ice for five minutes. Samples were diluted in 1 ml LB medium and incubated on the thermomixer for 45 minutes at 37 °C. After centrifugation for 5 min at 250 g supernatant was discarded and pelleted bacteria were resuspended in 200 µl fresh LB medium. Bacteria were plated on selective ampicillin (100 mg/ml) agar plates and incubated for 16 h at 37 °C.

#### 3.2.1.5 *DNA preparation (Mini-prep.)*

Single colonies grown on selective agar plates were selected and inoculated into 3 ml LB medium containing ampicillin (100 mg/ml). The cultures were incubated for 16 h at 37 °C at 200 rpm in an orbital shaker. 3 µl of bacterial culture was plated on selective agar plates; 2 ml culture was used to extract plasmid DNA.

Bacteria were pelleted by centrifugation for 10 min at 1500 g and 4 °C and resuspended in 100 µl E1 resuspension buffer (all “E” buffers were from JetStar 2.0 plasmid purification kit). By adding 100 µl E2 and inverting three times, cells were lysed by alkaline lysis for 5 min at room temperature. Adding 100 µl E3 buffer leads to precipitation of bacterial DNA and proteins. After centrifugation for 20 min at 16100 g and 4 °C, the supernatant was transferred into 900 µl ice-cold 100 % Ethanol. Plasmid DNA is precipitated by centrifugation for 30 min at 16100 g and 4 °C. After sequential wash with 500 µl 70 % and 100 % Ethanol the pellet was air-dried for 15 min and resuspended in 30 to 50 µl 1xTE.

#### 3.2.1.6 *Sequencing*

Samples were prepared from purified plasmid DNA in a concentration of 50 - 100 ng/µl in a total volume of 15 µl. For sequencing a general pJET1.2 primer was added by Eurofins Genomics, Germany.

Sequences obtained from Eurofins were compared to *in silico* sequences generated with MacVector by the ClustalW alignment tool. Plasmids with matching sequences were re-transformed into DH5α and inoculated into 400 ml LB medium. After overnight incubation at 37 °C and 200 rpm bacteria were harvested and plasmid DNA was extracted (see chapter 3.2.1.7).

#### 3.2.1.7 *High amount DNA preparation (Maxi-prep.)*

A 400 ml LB bacteria culture was incubated 16 h at 37 °C and 200 rpm in an orbital shaker. Plasmid DNA was extracted with the JetStar 2.0 plasmid purification kit according to manufacturer's instructions.

After centrifugation for 15 min at 2800 g and 4 °C, cells were resuspended in 10 ml E1 buffer. Lysis was done by adding 10 ml E2 buffer and incubating for 5 min at room temperature. For neutralization E3 buffer was added and after inverting the suspension a centrifugation for 30 min at 20000 g and 4 °C followed. Supernatant was filled into an equilibrated column. Bound DNA was washed with 60 ml E5 buffer and eluted from the column with 15 ml elution buffer E6 into a 50 ml tube. By adding 10.5 ml 2-propanol and centrifugation for 45 min and 2800 g plasmid DNA was precipitated. Pellet was washed with 70 % and 100 % Ethanol, air-dried for 15 min and resuspended in 500 to 1000 µl 1xTE buffer. DNA concentration was determined with NanoDrop® ND-1000 spectrometer and samples were stored at -20 °C.

#### 3.2.1.8 *Restriction digest and dephosphorylation*

For control digest of mini-plasmid preparations 2-5 units restriction enzyme for 1-2 µg DNA were used. Sample was diluted to 20 µl total volume in the appropriate buffer (NEB 1.1, 2.1, 3.1 or CutSmart). In restriction digests for further cloning (vector and insert digest) 4 µg plasmid was digested in higher volume. After incubation for 1 h at 37 °C enzymes were heat inactivated by incubating for 20 min at 65 °C or 80 °C according to manufacturer's instructions.

After heat inactivation of restriction digest, 1 unit CIP was added to the sample to prevent re-ligation of vector DNA. Directly after incubation for 30 min at 37 °C samples were loaded into an agarose gel.

#### 3.2.1.9 *Agarose gel electrophoresis*

DNA was separated according to expected fragment size in a 1-1.5 % agarose gel supplemented with 3 µl/100 ml ethidium bromide. 6xOrange G was added to the samples before loading them onto the gel. GeneRuler DNA Ladder Mix was used to determine fragment sizes. Electrophoresis was done at 100 V and 400 mA for 1 h.

Separated DNA fragments were either extracted for ligation and further cloning (chapter 3.2.1.10) or analyzed by UV excitation in a gel documentation system.

#### 3.2.1.10 *Gel extraction*

DNA fragments were cut from agarose gel and purified with the NucleoSpin Gel and PCR Clean-up kit according to manufacturer's instructions.

In brief, agarose containing DNA was dissolved in 200 µl NT buffer per 100 µg gel for 5-10 min at 50 °C. This sample was loaded onto a SpinColumn and centrifuged for 30 sec at 11000 g. After washing two times with 700 µl NT3 buffer and drying column membrane by centrifugation for 1 min at 11000 g, DNA was eluted twice with 15 µl water. Eluted DNA concentration was ascertained with NanoDrop® ND-1000 spectrometer.

#### 3.2.1.11 *Ligation*

Ligation was performed to covalently combine the digested insert with the digested vector fragment, both having compatible ends according to restrictions enzymes. Ligases were used as enzymes to build phosphodiester bounds between 3'-OH and neighboring 5'-phosphate groups. Two controls were conducted. Re-ligation control was performed without insert. Incomplete digest control for vector DNA digest was

conducted without the insert and the ligase. Ligation reaction mix was composed as following:

0.5 µl	T4 DNA ligase
1.5 µl	10x T4 DNA ligase buffer
x µl	Vector
y µl	Insert
1 µl	10 mM ATP
2 µl	50 % PEG 6000
ad 15 µl	H <sub>2</sub> O

The molar ratio of Vector:Insert was between 1:5 to 1:10, calculated from molecular weight and concentration of fragments. After mixing, reactions were incubated for 12-16 h at 16 °C. Transformation of ligated plasmids was done with 4 µl of ligation reaction in chemical competent DH5α cells.

#### 3.2.2 Cell culture

For all my experiments I used HEK293 cell lines. In Table 18 a short description of the cell line, the AGV internal identification and the information about cultivation conditions is listed.

##### 3.2.2.1 Cultivation of *HEK293* cell lines

HEK293EBNA1<sup>+</sup> cells were cultivated at 37 °C and 5 % CO<sub>2</sub> in DMEM medium supplemented with 8 % FBS, 220 µg/ml G418 and 1x Pen/Strep (100 units/ml Penicillin and 100 µg/ml Streptomycin). Cells were grown to 80 % confluence on 15 cm dishes and split 1:4 to 1:6 every three to four days. To split cells they were washed with 10 ml PBS and treated with 2 ml 0.25 % Trypsin-EDTA for 5 min at room temperature. After incubation cells were resuspended in new growth medium

and partly transferred to a new cell culture dish. For cell lines expressing additional sctetR fusion proteins, puromycin was added according to Table 18.

Table 18 Essential cell line information

Cell lines (AGV identification)	Description	Medium
<b>HEK293 (#43)</b>	Human embryonic kidney cells transformed with adenovirus 5	DMEM, 8 % FBS, 1 % Penicillin/Streptomycin
<b>HEK293EBNA1<sup>+</sup> (#1803)</b>	Human embryonic kidney cells, stably expressing EBNA1	DMEM, 8 % FBS, 1 % Penicillin/Streptomycin, 220 µg/ml G418
<b>HEK293EBNA1<sup>+</sup> + sctetR (#1456)</b>	Human embryonic kidney cells, stably expressing sctetR	DMEM, 8 % FBS, 1 % Penicillin/Streptomycin, 700 ng/ml Puromycin
<b>HEK293EBNA1<sup>+</sup> + sctetR:H3.3 (#2126)</b>	Human embryonic kidney cells, stably expressing sctetR:H3.3	DMEM, 8 % FBS, 1 % Penicillin/Streptomycin, 300 ng/ml Puromycin
<b>HEK293EBNA1<sup>+</sup> + sctetR:CENP-A (#2419)</b>	Human embryonic kidney cells, stably expressing sctetR:CENP-A	DMEM, 8 % FBS, 1 % Penicillin/Streptomycin, 300 µg/ml Puromycin
<b>HEK293EBNA1<sup>+</sup> + sctetR:CENP-A<sup>mut</sup> (#2506)</b>	Human embryonic kidney cells, stably expressing sctetR:CENP-A <sup>mut</sup>	DMEM, 8 % FBS, 1 % Penicillin/Streptomycin, 300 µg/ml Puromycin
<b>HEK293EBNA1<sup>+</sup> (#2901)</b>	Generated from cell line #43 by integrating expression plasmid 3279	DMEM, 8 % FBS, 1 % Penicillin/Streptomycin, 400 µg/ml G418

#### 3.2.2.2 Determination of cell number

To count the cells, they were trypsinized and resuspended in growth medium as described before. Cell suspension was transferred into a falcon tube and 10 µl were used to spot on a Neubauer cell counting chamber. Total cell number was determined by following equation:

$$total\ cell\ number = \frac{Cell\ count}{counted\ squares} \times 10^4\ cells/ml \times volume\ [ml]$$

Cells were seeded according to the calculated cell concentration after counting.



#### 3.2.2.3 *Generation of stable cell lines*

For generation of stable cell lines expressing EBNA1 from the CMV promoter, cells were seeded in a 6-well to a density of  $2 \times 10^5$  and transfected with 2  $\mu\text{g}$  linearized expression plasmid 3279 (see Table 14) using Lipofectamine2000 according manufacturer's instructions (detailed information about transfection procedure is given in 3.2.2.5). Transfected cells from one 6-well were plated in medium with 300, 400 and 500  $\mu\text{g}/\text{ml}$  G418 on three 15 cm culture dishes the next day. After two to three weeks, single colonies from the 400  $\mu\text{g}/\text{ml}$  G418 selected cells were separated and expanded in 6 well plates. Expression of EBNA1 was verified by Western Blot and immune fluorescent staining.

#### 3.2.2.4 *Cryoconservation*

HEK293EBNA1<sup>+</sup> cells were grown to confluence and three vials were frozen from one 15 cm cell culture dish. Cells were pelleted and resuspended in a suspension of 90 % FBS mixed with 10 % DMSO. They were frozen to  $-80^\circ\text{C}$  and after a few days cells were transferred to the liquid nitrogen tank for long-term storage.

For thawing, cells were quickly warmed to  $37^\circ\text{C}$ , washed with 10 ml fresh medium to remove remaining DMSO and plated on a new 15 cm cell culture dish in 20 ml of fresh growth medium.

#### 3.2.2.5 *Transfection for plasmid maintenance experiments*

Transfections for plasmid maintenance experiments with Lipofectamine 2000 were performed according to manufacturer's instructions. In brief,  $4 \times 10^5$  cells were seeded per well into a 6-well plate and transfected 24 h later. 2  $\mu\text{g}$  of plasmid DNA and 4  $\mu\text{l}$  of Lipofectamine2000 were diluted separately with Opti-MEM medium to a final

volume of 50  $\mu$ l each and incubated for five minutes. Subsequently, both solutions were mixed and incubated for 20 min at room temperature. The resulting 100  $\mu$ l transfection solution was applied to the cells drop-wise. 24 hours after transfection cells were transferred to a 15 cm cell culture dish in fresh DMEM growth medium containing 120  $\mu$ g/ml hygromycin.

#### 3.2.2.6 *Transient transfection*

Transient transfections were carried out with the transfection reagent polyethylenimin (PEI) in 6-well plates.  $4 \times 10^5$  cells were seeded per well into a 6-well plate and transfected 24 h later. 2  $\mu$ g DNA of corresponding plasmid were mixed with 300  $\mu$ l DMEM<sup>-FBS</sup> (DMEM without FBS but supplemented with Penicillin/Streptomycin and 1 % HEPES). PEI transfection reagent was also mixed with 300  $\mu$ l DMEM<sup>-FBS</sup> in a ratio of 3  $\mu$ g PEI per 1  $\mu$ g DNA. Both solutions were combined and vortexed for 10 sec. Reaction was incubated for 20 min at room temperature and drop-wise distributed onto the cells. After four hours incubation, the medium was replaced by normal cell growth medium and cells were incubated for 16 h at 37 °C in the incubator.

The next day, cells were plated on 10 cm (1 well) or 15 cm (2 wells) for further experiments.

#### 3.2.2.7 *Cell synchronization with thymidine block*

Synchronization was done by a double thymidine block and release. Thymidine blocks cells during S-phase by influencing the formation of a regulatory dCTP pool, which is important for DNA synthesis (Bjursell and Reichard, 1973).

Cells were seeded in 6-wells,  $2 \times 10^5$  cells per well, in 1 ml normal growth medium and incubated for 8 h at 37 °C and 5 % CO<sub>2</sub>. 1 ml thymidine medium (8 mM

thymidine in normal DMEM growth medium) was added and incubated for 16 h. Cells were released from block by washing twice with PBS and coating them with 1 ml growth medium. Second block was done 8 h later by adding again 1 ml 8 mM thymidine medium for 16 h. Final release followed by washing twice with PBS and coating cells with 2 ml growth medium for 8 h, until mitosis.

#### 3.2.2.8 *Transient transfection of synchronized cells*

Cells from three 6-well plates were released from block as described in chapter 3.2.2.7. After 8 h incubation at 37 °C and 5 % CO<sub>2</sub>, cells were transfected according to PEI transfection protocol (described in detail in chapter 3.2.2.6) with three different reporter plasmids, the tetO-DS sc tetR:CENP-A reporter plasmid, the tetO-DS sc tetR:HMGA1a control and the *oriP* control.

I replaced normal growth medium by 1 ml DMEM without serum. One 6-well plate was transfected with one reporter plasmid. To each well 600 µl transfection reaction mix, containing 200 ng reporter plasmid and 1.8 µg fill-up plasmid (LacO plasmid), was added drop-wise after incubation of transfection mixture for 20 min at room temperature. Cells were incubated for 4 h at 37°C and 5 % CO<sub>2</sub>. Afterwards medium was replaced by normal growth medium containing FBS and placed into the incubator. One well each also contained also 2 µg/ml doxycycline (day0). On the next day wells were transferred to 10 cm cell culture dishes. Doxycycline was added to one additional plate as well (day1). On day2 and day4 medium of one plate each was replaced by medium containing 2 µg/ml doxycycline. Cells were split every three to four days and doxycycline medium was replaced every second day. On day9 cells were harvested for FACS analysis (see chapter 3.2.4).

#### 3.2.2.9 *Transient transfection for live cell imaging*

In total 5 µg plasmids were transfected with Xfect™ Polymer as transfection reagent according to manufacturer's instructions. A short description about procedure is given in the following.  $4 \times 10^5$  HEK293EBNA1<sup>+</sup> cells were seeded in 6-wells the day before transfection. Reporter plasmid carrying scetR:GFP:CENP-A, tetO-DS and the Cas9 targeting sites (1.5 µg) (6359, Table 14) was mixed with CMV-Cas9:3xmCherry (0.5 µg) and the sgRNA expression plasmid (3 µg) in 100 µl Xfect™ reaction buffer total volume. This mixture was vortexed for 5 sec. Afterwards 1.5 µl Xfect™ Polymer was added to the diluted plasmids and vortexed again for 10 sec. After incubation for 10 min at room temperature and brief centrifugation, the 100 µl transfection solution was added drop-wise to the cell culture medium. Cells were incubated overnight at 37 °C in the incubator. The next day cells were transferred on 10 cm dishes and grown for additional three days. On day four after transfection cells were counted and seeded into ibidi® live cell imaging slides. Depending on the experiment, cells were stained with live cell tubulin dye, SiR-tubulin (for further details see chapter 3.2.10).

#### 3.2.3 **Plasmid rescue assay**

For the analysis of long-term plasmid maintenance reporter plasmids were transfected and selected for two weeks. After two weeks selection pressure was either removed or cells were kept further under selection as controls. Every two weeks a plasmid rescue assay was performed. Therefore, cells were lysed and plasmid DNA was enriched according to the HIRT protocol (Hirt, 1966).

In detail, after washing confluent cells with 10 ml PBS on the dish, cells were equilibrated in 5 ml TEN buffer. TEN was removed and cells were coated with 1.5 ml TEN buffer and an equal volume of 2xHIRT buffer was added for cell lysis. Cell lysis was performed on the dish by tilting the plate for 1 min. The lysate was collected in a

flip-cap tube and genomic DNA and proteins precipitated at 4 °C for 16 h, in the presence of 1 M NaCl. After centrifugation for 45 min at 4 °C and 20000 g, DNA was purified by phenol-chloroform extraction, precipitated with ethanol and resuspended in 100 µl 1xTE buffer.

DNA samples were digested with 40 units DpnI in presence of 0.5 µl RNase. 350 ng digested DNA was electroporated into Electromax DH10β competent cells. For electroporation all buffers and cuvettes were cooled throughout the procedure. 100 µl of bacteria were diluted with 500 µl ice-cold water and 100 µl of this diluted suspension were mixed with 50 µl DNA (350 ng) sample. Electroporation was performed in 1 mm cuvettes at 25 µF and 2.5 kV. After electroporation cells were transferred in 3 ml of LB medium and incubated on a shaker at 37 °C for 45 min. Afterwards cells were pelleted, plated on selective ampicillin agar plates and incubated overnight at 37 °C. The ampicillin-resistant colonies, representing the number of recovered plasmids, were counted the next day. For calculations the colony number obtained at the first plasmid rescue, one week after removal of selection pressure, was used for normalization.

#### 3.2.4 Flow cytometry

Cells for flow cytometry were harvested by trypsinization and resuspending in 5 ml PBS containing 2 % FBS. After centrifugation at 250 g and 4 °C for 7 min cell pellet was resuspended in 1 ml PBS and transferred into FACS tubes. These tubes were kept on ice during measurement time. Cells were diluted that only 1000 to 1500 events per second were measured in BD FACSCalibur device. Measurement stopped after counting 10<sup>5</sup> events. Gating for living, intact cells was done afterwards by analysis in the FlowJo program. Therefore, only cells having defined size (forward scatter) and defined granulation profile (side scatter) were gated. Out of this population the percentage of gfp positive cells was determined.

#### 3.2.5 Western Blot

RIPA protein extraction was performed by harvesting HEK293 cells with trypsin and transferring into falcon tubes. After washing twice with ice-cold PBS cell suspension was centrifuged (600 g, 7 min, 4 °C). Pellet was resuspended in two pellet volumes RIPA extract buffer supplemented with 1x complete protease inhibitor and incubated for 20 minutes on ice. Lysate was vortexed for 15 sec and centrifuged at 16100 g, 20 min, 4 °C. The protein concentration of the supernatant was determined using Bradford reagent and 20 to 50 µg protein extract with 1x Laemmli buffer was loaded on the SDS gel after boiling for 5 min at 95 °C. The remaining supernatant was stored at -20 °C for long time storage.

Depending on predicted proteins size extracts were separated on an 8 % or 13 % SDS-polyacrylamide gel and blotted on Amersham Hybond ECL membrane using semiDry blotting system. Protein separation and transfer onto the membrane was verified by ponceau stain and the membrane was blocked for 1 h in 5 % milk in PBS, 0.1 % Tween-20. Incubation with primary antibody (dilution see Table 5), was performed in 2.5 % milk in PBS, 0.1 % Tween-20 for 16 h at 4 °C. Afterwards, membrane was washed 3x 5 min with PBS, 0.1 % Tween-20 and incubated with secondary HRP coupled antibody 1:10000 in 2.5 % milk in PBS, 0.1 % Tween-20 for 1 h at room temperature. After repeated washing steps (3x 10 min in PBS, 0.1 % Tween-20), revelation was done using ECL containing H<sub>2</sub>O<sub>2</sub> on CEA Blue Sensitive X-ray films.

#### 3.2.6 Co-Immune precipitation of sctetR fusion proteins

For co-immune precipitation 5x 10<sup>7</sup> cells were harvested. Cells were trypsinized and resuspended in ice-cold PBS supplemented with 5 % FBS. After centrifugation (250 g, 4 °C, 7 min) cell pellet was washed with 50 ml PBS and transferred into an eppendorf tube with 1 ml PBS. Cells were permeabilized by lysolecithin. Therefore, cell pellet

was resuspended in 1 ml permeabilizing buffer and 100 µl pre-warmed 1 % lysolecithin solution was added. Suspension was incubated for 5 min at 37 °C and 70 units MNase were added and mixed by inverting 5 times. Samples were incubated for 30 min at 37 °C. MNase reaction was stopped by adding 50 µl 0.5 M EGTA (final concentration 20 mM). Nuclei were pelleted by centrifugation (400 g, 3 min, 4 °C) and was performed in 500 µl lysis buffer on ice for 30 min. By high speed centrifugation at 16100 g for 15 min at 4 °C, cell fragments and aggregates were precipitated. Supernatant was used for co-immune precipitation. Two 50 µl aliquots were taken for MNase digest analysis and as input control for Western Blot. MNase digest control was treated with proteinase K and RNase A and DNA was extracted by phenol-chloroform extraction (as described in chapter 3.2.9.4). The supernatant for IP was transferred into 15 ml tubes and diluted with dilution buffer to a final concentration of 150 mM NaCl. IP buffer was supplemented with 0.5 % Sarcosyl, 0.1 % DOC and 0.8 % Triton X-100. After pre-clearing with protein G beads (3x washed with PBS) for 2 h, 100 µl tetR antibody coupled beads (see appendix for coupling procedure) were added and incubated overnight at 4 °C on orbital shaker. A sequential wash with 1 ml RIPA-300, RIPA-150 and 1xTE followed. Elution was performed with 2x50 µl of 1xTE+1 % SDS for 5 min at room temperature on an orbital shaker. Elutions were combined and supplemented with 5xLaemmli buffer. In an 8 % PAGE IP of tetR was approved, in a 13 % PAGE co-precipitation of histone proteins was investigated.

#### 3.2.7 Chromatin Immune Precipitation (ChIP)

For chromatin immune precipitation HEK293EBNA<sup>+</sup> cells were transfected with the tetO-DS sctetR:HMGA1a (3448) and sctetR:CENP-A (6083) reporter plasmids in presence and absence of 2 µg/ml doxycycline were used.

#### 3.2.7.1 *Cross-link*

Cells were trypsinized, washed twice in 10 ml PBS with and without 4 µg/ml doxycycline and after centrifugation (250 g, 4 °C, 7 min) cell pellet was resuspended in 10 ml PBS with and without 4 µg/ml doxycycline. An equal volume of PBS supplemented 2 % methanol-free formaldehyde was added and cells were fixed for 5 min on a roller at room temperature. The cross-link reaction was quenched with glycine (1.25 M). After incubation for one minute on the roller and 5 min on ice, cells were washed with once with ice-cold PBS. Nuclei preparation was performed in 10 ml ice-cold PBS, 0.5 % NP-40 for 10 min on ice. After pelleting, nuclei were resuspended in PBS containing 10% glycerol, pelleted as  $2 \times 10^7$  cell aliquots and snap frozen in liquid nitrogen.

#### 3.2.7.2 *Sonication*

Cross-linked cell pellets were thawed on ice and resuspended in 1 ml LB3+ buffer to a final concentration of  $2 \times 10^7$  cells/ml. Sonication was performed in AFA Fiber & Cap tubes (12x12 mm) at an average temperature of 5 °C to 7 °C. For sonication the Covaris S220 was used. Settings were established for HEK293EBNA1<sup>+</sup> cells in a concentration of  $2 \times 10^7$  cells/ml at 100 W, 150 cycles/burst, 10 % duty factor for 10 min.

#### 3.2.7.3 *Immune precipitation*

After sonication, chromatin concentration was measured by NanoDrop® ND-1000 Spectrometer and aliquots of 250 µg chromatin were prepared. Sheared chromatin was pre-cleared with 50 µl protein G beads (washed 3x in PBS, 50 % bead slurry prepared) for 2 h. 250 µg chromatin was incubated for 16 h at 4 °C with 50 µl of the 31B3 α-tetR monoclonal antibody supernatant and as IgG control 50 µl α-HA 12CA5



monoclonal antibody supernatant was used. BSA-blocked protein G beads (incubated beads 16 h on roller in blocking solution at 4°C) were added (50 µl/ 250 µg chromatin) and incubated for 4 h on orbital shaker at 4 °C. Sequential washing steps with LiCl, RIPA-300, RIPA-150 buffer and finally twice in 1xTE (pH 8.0) buffer were performed. Immuno-precipitated chromatin fragments were eluted from the beads by shaking twice at 1200 rpm for 10 min at 65°C with 100µl of TE, 1 % SDS. The elution was treated with 80 µg RNase A for 2 h at 37 °C and with 8 µg proteinase K at 65 °C for 16 h. DNA was purified using the NucleoSpin Gel and PCR Clean-up kit (and NTB binding buffer for SDS containing samples) according to manufacturer's instructions. Quantitative PCR analysis was performed as described in 3.2.8 and quantitative PCR values were represented as fold enrichment relative to isotype IgG control or % input calculated relative to an input standard curve. Chromatin sizes were verified by loading 1-2 µg eluted DNA on a 1.5 % agarose gel.

#### 3.2.8 Quantitative PCR (qPCR)

Quantitative PCR was performed with the Roche LightCycler® 480 System and the Light-Cycler®Fast-Start-DNA-Master-SYBR-Green-I (2xSYBR). 2 µl of ChIP elution were mixed with 8 µl master mix containing 5 µl 2xSYBR, 2.5 µl H<sub>2</sub>O, 0.5 µl 5 µM primer mix. Amplification was performed using the Roche SYBR standard program depicted in Table 19. qPCR primers are listed in Table 20.

Table 19 Roche SYBR Green qPCR standard program

	Temperature [°C]	Duration [s]	Cycles	Detection
<b>Pre-incubation</b>	95	300	1	
<b>Amplification</b>	95	1	45	
	60	10		
	72	10		
	75	3		single
<b>Melting curve</b>	97	1	1	
	67	10		
	97	(heat to 97°C)		continuous
<b>Cooling</b>	37	15		

Primer pairs, listed in Table 20, detected specific fragments on plasmid DNA. These fragments resulted from sonication and precipitation in ChIP experiments or resulted from restriction enzyme digest during STREP pull down protocol.

Table 20 qPCR primer

Primer (Schepers group internal identification)	Sequence [5' → 3']
DS_for (284)	TGTCATAGCACAATGCCACCAC
DS_rev (285)	GGTCAGGATTCCACGAGGGTAG
FR_for (276)	CGTGCTCTCAGCGACCTCG
FR_rev (277)	TCAAACCACTTGCCACAAAAAC
Sc4_for (280)	TCGGCGTCCACTCTCTTTCC
Sc4_rev (281)	CAGTAAGGTGTATGTGAGGTGCTCG
tetO_for (383)	GGGGGTGTTAGAGACAACCAGTG
tetO_rev (384)	GGCAGGGACCAAGACAGGTG
reference_for (756)	CGGCAACATCCTGGGGC
reference_rev (757)	CTGCTGGTAGTGGTCGGCG
AseI_for (758)	CACCACGATGCCTGTAGCAATG
AseI_rev (759)	AGCAATAAACCAGCCAGCCG
AflIII_for (764)	CAGTCAGAGAACCCCTTTGTGTTTG
AflIII_rev (765)	GACCACTAACCTTCGCTCCATACC
Cas9_for (791)	TTACTCTCTTCCCAAAGGATGTGC
Cas9_rev (792)	AAACCTGTCGTGCCAGAACTTG

Digest efficiencies were calculated according to Hagege et al. (Hagege et al., 2007) as following:

$$\% \text{ restriction} = 100 - \frac{100}{2[(Ct(res)-Ct(contr))_{dig} - (Ct(res)-Ct(contr))_{und}]}$$

#### 3.2.9 STREP pull down with Cas9:mCherry:TAP

HEK293EBNA1<sup>+</sup> cells were transfected with 2 µg reporter plasmids tetO-DS + sctetR:CENP-A either with or without Cas9 targeting sites and selected for two weeks with 120 µg/ml hygromycin. These cells were seeded again and transfected with the sgRNA expression plasmid (1 µg) together with the Cas9:mCherry:TAP

expression plasmid (1  $\mu$ g). After four days  $10^8$  cells were harvested by trypsinization and resuspension in PBS+5 %FBS.

#### 3.2.9.1 *Cross-link*

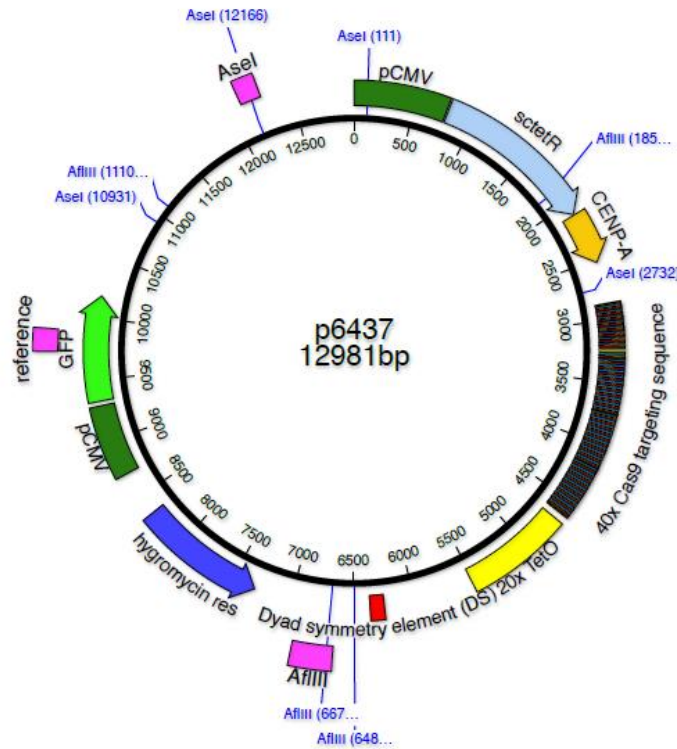
Cells were washed 2x with 20 ml PBS. After second wash cell pellets were resuspended in 10 ml PBS and equal volume of PBS supplemented with 2 % formaldehyde was added. Cross-link was stopped after 5 min incubation at room temperature on the roller by adding 2.2 ml 1.25 M glycine and incubation for 5 min at room temperature. Cells were washed in 10 ml ice-cold PBS. Nuclei preparation followed by 30 min incubation in 20 ml PBS + 0.5 % NP-40 on ice. Nuclei were centrifuged at 600 g for 10 min and 4 °C. After washing in ice-cold PBS, cells were resuspended in 5 ml 3.1 buffer (concentration  $2 \times 10^7$  cells/ml).

#### 3.2.9.2 *French Press*

To re-solubilize proteins and chromatin of interest, the 5 ml cell suspension in 3.1 buffer was loaded into the French Press cell. Suspension was stored on ice and pressed three times with a pressure of 13000 psi in a pre-cooled 40000 K cell. After applying high pressure, lysate was aliquoted into 1 ml aliquots and frozen at -80 °C. Two aliquots were directly used for STREP pull down.

#### 3.2.9.3 Restriction digest

Two different restriction enzymes were selected to fragment plasmids for STREP pull down. The combination of *AflIII* and *AseI* (Figure 18) led to a higher probability for separation of exactly the fragment of interest, tetO array + Cas9 targeting sequence.



**Figure 18 Representation of tetO-DS scetR:CENP-A + Cas9 plasmid**

tetO-DS scetR:CENP-A reporter plasmid with Cas9 targeting sites. *AseI* and *AflIII* restrictions sites are also depicted, together with qPCR products covering restriction sites.

For the digest, 120 units *AflIII* and 120 units *AseI* were added to  $2 \times 10^7$  cells. The 1 ml aliquots were distributed into 5 aliquots and incubated for 2 h at 37 °C with 1200 rpm. After incubation, aliquots were combined into one tube and centrifuged for 5 min at 500 g and 4 °C. 50 µl Controls were taken at each step to control resolubilization, digest efficiency and IP success by qPCR and Western Blot.

#### 3.2.9.4 *Pull down*

Supernatant obtained after restriction digest was supplemented to final concentrations of 150 mM NaCl, 0.5 % Sarcosyl, 0.1 % DOC, 0.5 % Triton X-100 and 1x complete protease inhibitor with these stock solutions. Pre-clearing was performed by adding 100 µl protein G sepharose beads (3x washed with PBS) for 1 h at 4 °C. Beads were removed by centrifugation for 5 min at 300 g and 4 °C. Supernatant was incubated with 150 µl STREP-tactin® bead slurry (equilibrated 3x with 3.1 buffer) for 16 h at 4 °C by constantly rotating on orbital shaker. Beads were sequentially washed with RIPA-300 buffer, RIPA-150 buffer and 1x TE buffer. Elution was performed by incubating the beads two times for 5 min at room temperature with 100 µl 1x TE + 0.5 % SDS. For Western Blot analysis, the cross-link was reverted for 1 h at 65 °C, 5x Laemmli was added and samples were boiled for 5 min at 95 °C.

qPCR controls were incubated for 1 h with RNase A at 37 °C and subsequently for 16 h with proteinase K at 65 °C in 400 µl total volume. DNA was extracted by adding 400 µl Phenol and incubation for 5 min at room temperature on orbital shaker. After centrifugation at 16100 g for 20 min and 4 °C, water-phase was transferred into new tube. 400 µl Chloroform-Isoamylalcohol was added and incubated for 5 min on the orbital shaker. This step was repeated after centrifugation at 16100 g, 15 min, 4 °C. Water-phase was transferred into new tube and 2.5 volumes (1000 µl) 100 % ethanol and 40 µl Sodium-acetate were added. DNA precipitated overnight at -20 °C. Pelleting DNA was done by centrifugation at 16100 g, 30 min, 4 °C. After washing DNA pellet with 70 % and 100 % ethanol, it was resuspended in 50 µl 1xTE buffer. DNA concentration was determined using the NanoDrop® ND-1000 Spectrometer. Digestion and re-solubilization efficiency was investigated by loading 10 µl samples into a 1.5 % agarose gel and qPCR was performed as described in chapter 3.2.8.

#### 3.2.10 Live cell imaging with Cas9:3xmCherry

HEK293EBNA1<sup>+</sup> cells were transfected with pCON<sup>CENP-A</sup> plasmids containing Cas9 targeting sites and co-transfected with the sgRNA and Cas9:3xmCherry expression plasmids as described in chapter 3.2.2.9. Four days after transfection cells were trypsinized and collected in a 15 ml tube. Cells were counted and 300 µl suspension were seeded into ibidi® live cell imaging slides in a concentration of 5 to 8x 10<sup>4</sup> cells. Cells settled for 6 h at 37 °C in the incubator. 50 µl medium was removed and replaced with 50 µl medium supplemented with SiR-tubulin to a final concentration 100 nM. This concentration was constantly maintained during imaging.

#### 3.2.11 Immune fluorescence

HEK293EBNA1<sup>ΔGA+</sup> cells were transfected with pCON<sup>CENP-A</sup> plasmids expressing the sctetR:GFP:CENP-A fusion protein and incubated for four days in the incubator.

##### 3.2.11.1 *CENP-C/Ndc80 and EBNA1 co-staining*

Four days after transfection 2x 10<sup>5</sup> cells were seeded onto Poly-L-Lysine coated cover slips. (Cover slips were coated by washing with 80 % ethanol and incubating with 2 ml 0.01 % Poly-L-Lysine solution for 20 min at room temperature in 6-wells. After incubation Poly-L-Lysine solution was removed and cover slips were air dried for 20 min.) Seeded cells were settled for 8 h at 37 °C in the incubator in 1 ml growth medium. 1 ml 8 mM thymidine containing growth medium was added and cells were blocked for 16 h in the incubator. By washing 2x with PBS and adding 2 ml growth medium the thymidine block was released. 8 h after release cells were fixed in mitosis. Cells were fixed for 8 min in 4 % formaldehyde in PBS + 0.1 % Triton X-100. After washing with PBS + 0.1 % Triton for 5 min at room temperature, cells were blocked with PBS + 0.1 % Triton containing 5 % FBS for 1 h at room temperature.

Subsequently, 30  $\mu$ l of the first primary antibody in PBS + 0.1 % Triton containing 2.5 % FBS, CENP-C or Ndc80, was dropped onto the cells (concentration see Table 5) and incubated overnight at 4 °C. After washing 3x 5 min with PBS + 0.1 % Triton, the incubation with 30  $\mu$ l secondary antibody,  $\alpha$ -mouse Cy5 1:100 in PBS + 0.1 % Triton containing 2.5 % FBS, for 1 h at room temperature was performed. From this step of protocol all further steps were performed in the dark. After washing 3x 5 min with PBS + 0.1 % Triton, cells were incubated with 30  $\mu$ l EBNA1 antibody, 1:10 diluted in PBS + 0.1 % Triton containing 2.5 % FBS for 1 h at room temperature. Subsequently, cells were washed 3x 5 min and the secondary antibody,  $\alpha$ -rat Cy3 1:100, was added and incubated for 1 h at room temperature. Counter staining of DNA was done with 50  $\mu$ l of 250 ng/ml DAPI solution for 2 min after washing 3x 5 min. After a final wash step in PBS + 0.1 % Triton, cells were mounted in vecta shield on a microscope slide and sealed with nail polish. Slides were stored at 4 °C in the dark until imaging.

#### 3.2.11.2 *EBNA1 staining*

HEK293EBNA1<sup>ΔGA+</sup> cells were co-transfected with pCON<sup>CENP-A</sup> plasmids expressing the sctetR:GFP:CENP-A fusion protein and a RFP:CENP-A expression plasmid. Two and four days after transfection, 4x 10<sup>5</sup> cells were seeded on Poly-L-Lysine covered cover slips. Fixation and blocking was done as described in chapter 3.2.11.1. After blocking 30  $\mu$ l EBNA1 antibody, 1:10 diluted in PBS + 0.1 % Triton containing 2.5 % FBS for 1 h at room temperature, and after washing 3x 5 min the secondary antibody,  $\alpha$ -rat Alexa647 1:100, was added and incubated for 1 h at room temperature in the dark. Counter staining of DNA with DAPI and mounting cells onto microscope slides was performed as described above.

#### 3.2.12 Microscopy

For microscopy of fixed samples I used the confocal laser scanning microscope (CLSM), Leica TCS SP2.

For live cell imaging I used a fluorescence microscope without confocal scanning. I collaborated with Steffen Dietzel and Andreas Thomae of the Biomedical center in Munich (BMC) and used the Leica inverted motorized live cell fluorescence microscope, Leica DMI8.

##### 3.2.12.1 *Fixed samples*

Fixed samples were imaged at the confocal laser scanning microscope (CLSM), Leica TCS SP2 with a 63x oil objective. The pinhole was set to 0.9 AU and a 63x oil objective with a numerical aperture of 1.4 was used. Argon laser intensity was set to 30 % and DPSS 501 nm, HeNe 594 nm and HeNe 633 nm were also used. Excitation intensity of DAPI was kept below 10 % UV to avoid bleaching of other fluorescent dyes. GFP was excited with 488 nm (~ 15 %), mCherry and Cy3 were excited with 550 nm (~ 25 %) and Cy5 or Alexa647 was excited with 633 nm (~ 25 %). Gain of PMTs was set between 700 and 1000 V as recommended in Leica TCS SP2 instructions. Imaging was done in stacks of around 15 to 20  $\mu\text{m}$  height with a z-step size of 500 nm. The pixel size was around 100 nm and a frame average of 3 frames was performed for each color to improve signal to noise ratio.

##### 3.2.12.2 *Live cell Imaging*

Live cell imaging was performed at the BMC at the Leica inverted motorized live cell fluorescence microscope, Leica DMI8. I used a 40x dry objective with a numerical aperture of 0.6. Microscope stage was surrounded by dark chamber and was heated



to 37 °C and the CO<sub>2</sub> concentration was set to 6 %. The LED quad filter SpX-Q was used for excitation and emission detection. Excitation wavelengths generated by the spectrax light engine® (Lumencor® Inc., USA) and filtered by the LED quad were 390/22 nm, 470/24 nm, 550/15 nm and 640/30 nm. For imaging settings were used as shown in Table 21.

Table 21 Settings for live cell imaging

Signal	Excitation wavelength	illumination intensity	illumination time	Bin	Gain
<b>bright field</b>	BF lamp	30	100 ms	3x3	1.0
<b>gfp</b>	470 nm (7 %)	100 %	200 ms	3x3	5.0
<b>rfp</b>	550 nm (6 %)	100 %	200 ms	3x3	5.0
<b>SiR-tubulin</b>	640 nm (4 %)	100 %	200 ms	3x3	2.0

In each experiment I chose ten positions to image every 4 minutes. At each position a stack of 7.5 µm with a z-step size of 1.5 µm was recorded. Total imaging time was set to 24 h.

## 4 RESULTS

Human centromeres are functionally defined regions on chromosomes that regulate kinetochore assembly during mitosis. It is known that centromere identity is not genetically determined. Centromere positioning is rather specified epigenetically by CENP-A, the centromere specific H3 histone variant (Black et al., 2007; Mendiburo et al., 2011; Warburton et al., 1997). However, little is known about other potential epigenetic modifications contributing to centromere inheritance.

We demonstrated, in collaboration with Patrick Heun's group, that artificial targeting of CENP-A to plasmids is sufficient to establish neo-centromere activity on plasmids in *Drosophila* Schneider S2 cells (Mendiburo et al., 2011). With the pCON<sup>CENP-A</sup> vector, we transferred the CENP-A targeting system into human cells and I investigated centromere establishment and maturation.

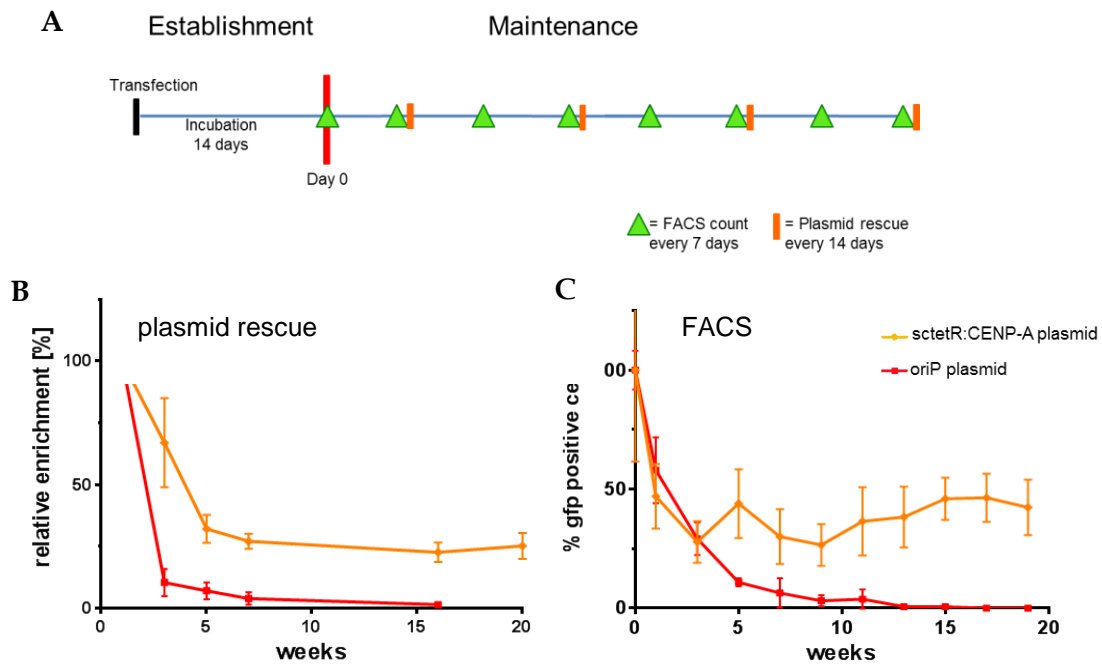
Detailed information about the pCON<sup>CENP-A</sup> plasmid is given in chapter 1.8. In brief, the pCON<sup>CENP-A</sup> contains the replication element (DS) of EBV's latent replication origin (*oriP*) and a tetO array as *cis*-acting elements. Plasmid replication is initiated by EBNA1 binding to DS. Segregation is mediated by CENP-A targeting to tetO thus inducing centromere activity. *oriP* plasmids served as reference in the plasmid maintenance analyses.

In order to investigate the stability of centromeres established by targeting of CENP-A and the resulting epigenetic inheritance of these centromeres, we first examined the plasmid maintenance for more than 20 weeks. Plasmids were transfected into HEK293 cells and plasmid abundance was analyzed weekly by measuring gfp-reporter gene expression and plasmid rescue experiments. In the following, I will describe the experimental setup, the observations of long-term plasmid maintenance and their functional relevance.

### 4.1 Long term pCON<sup>CENP-A</sup> maintenance

We assumed, that the establishment of an inheritable centromere identity leads to stable maintenance of pCON<sup>CENP-A</sup> over many cell generations independent of selective pressure. In previous experiments, a stable and inheritable centromere was established at the pCON<sup>CENP-A</sup> plasmid after targeting CENP-A. These plasmids were maintained for three or four weeks under selective pressure (Figure 14 and Figure 15). In order to investigate long-term maintenance, we decided to follow plasmid abundance in HEK293 cells after an establishment time of two weeks in selective medium for around 20 weeks without selection.

Therefore, Lara Schneider transfected HEK293EBNA1<sup>+</sup> cells with *oriP* and pCON<sup>CENP-A</sup>. After the initial establishment time the selection medium was replaced by normal cell culture growth medium. Every seven days, plasmid maintenance was monitored by FACS (see Figure 19 A). In FACS analysis (fluorescence activated cell sorting/scanning) cells are scanned according to their size, granulation level and fluorescence intensity. Living cells were defined according to size and granulation of reference untransfected HEK293 cells. Within this cell population, the proportion of gfp<sup>+</sup> expressing cells was determined. Every second week, the plasmid abundance was also measured more directly by plasmid rescue assays. In this experiment, cells are lysed and low molecular weight DNA is extracted. To discriminate between bacterial “input” DNA that was initially transfected on the one hand and replicated and segregated DNA on the other, a *DpnI* digest was performed. The transfected DNA that was initially purified from bacteria is *dam* methylated. *DpnI* is a *dam* methylation sensitive enzyme that digests DNA that carries a methylated A in the GATC-motif. After *DpnI* digest, only replicated plasmid DNA is intact and can be re-transformed into bacteria to represent the amount of plasmids that were present in the cells.



**Figure 19: Plasmid rescue and FACS analysis of plasmid maintenance**

A) Timeline of experimental setup. Establishment of plasmids is under selection pressure for 14 days. After two weeks selection is removed and plasmid maintenance is analyzed every two weeks by plasmid rescue and weekly FACS counts.

B) pCON<sup>CENP-A</sup> plasmids are maintained stable over 5 months. *oriP* is lost already 5 weeks after selection removal.

Data: Lara Schneider (n=4, Mean+SD)

*oriP* plasmids were lost within three to five weeks after removal of selective pressure (Figure 19 B, red line). A similar plasmid loss rate was observed by measuring the *gfp*<sup>+</sup> level, expressed as destabilized GFP on the reporter plasmids (Figure 19 C, red line). In contrast, the pCON<sup>CENP-A</sup> plasmid was stably maintained after a copy number stabilization phase at a level between 30 and 50 % relative to the starting amount (Figure 19 B+C, yellow line).

In conclusion, targeting of CENP-A to the tetO array on pCON<sup>CENP-A</sup> leads to stable plasmid maintenance for a long time frame over 5 months. It is very likely that an inheritable centromere identity is established on the plasmid within two weeks establishment time under selective conditions.

Long-term stability is an appropriate prerequisite for the use of pCON<sup>CENP-A</sup> as gene therapy vector. However, before developing clinical applications with pCON<sup>CENP-A</sup> it is essential to functionally characterize it.

As a first step, I decided to investigate the targeting of sctetR:CENP-A to the tetO array and its segregation mechanism.

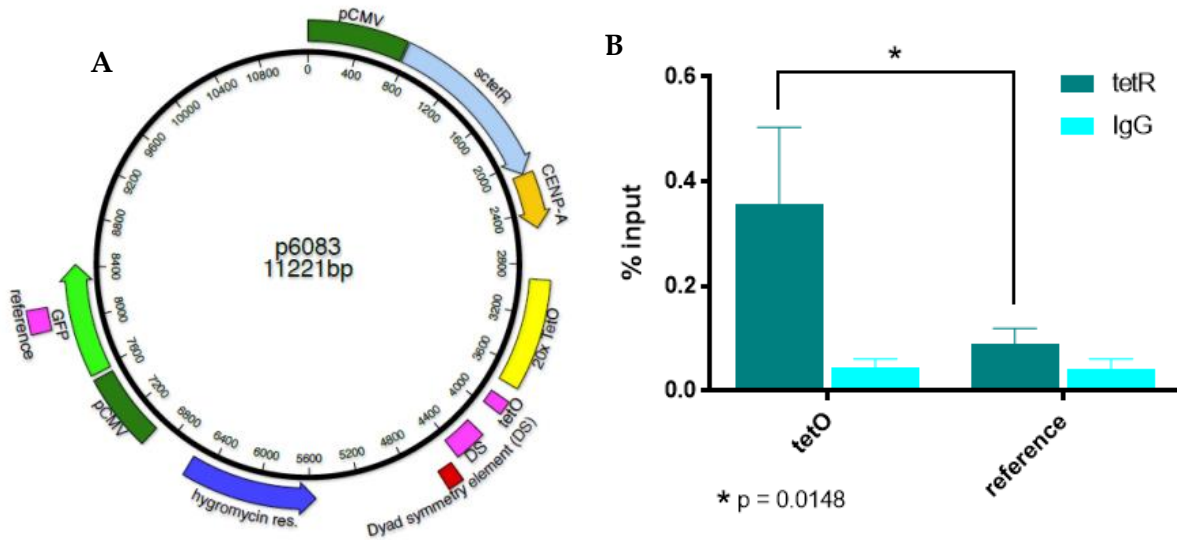
### 4.2 sctetR:CENP-A is site specifically targeted to tetO sites

The artificially induced neo-centromere is established after targeting sctetR:CENP-A to a tetO array on a plasmid system. Besides the tetO, the reporter plasmid harbors the EBV latent replication origin DS. To address the question if sctetR:CENP-A is site specific targeted to the tetO array on pCON, two different approaches were used. A ChIP experiment verified the sctetR:CENP-A fusion protein binding to tetO. Immune fluorescence was employed to study the co-localization between sctetR:GFP:CENP-A and the plasmids by overlapping signals resulting from confocal microscopy and its resolution limitations.

#### 4.2.1 ChIP directly demonstrates interaction with tetO

For the functional characterization of pCON<sup>CENP-A</sup> it was important to know if sctetR:CENP-A is site-specifically targeted to the tetO array on the plasmids. The binding of sctetR:CENP-A to tetO was determined in a ChIP experiment, since ChIP is a sensitive method to identify DNA-protein interactions. For ChIP the chromatin and chromatinized plasmids were fixed with formaldehyde and fragmented by sonication. After immune precipitation using a tetR-specific antibody, the tetR protein interacting DNA fragments were quantified by qPCR. The enriched DNA fragments at the region next to the tetO repeats and a reference region that is located approximately 5 kb apart from tetO were compared.

To perform a tetR specific ChIP, I first had to generate a suitable  $\alpha$ -tetR antibody in collaboration with the Monoclonal Antibody Core Facility of the Helmholtz Center Munich. Detailed information about antibody validation is given in the appendix.



**Figure 20: ChIP with  $\alpha$ -tetR monoclonal antibody**

A) Representation of tetO-DS scetR:CENP-A reporter plasmid with primer binding sites for qPCR analysis indicated in magenta (tetO, DS and reference region)

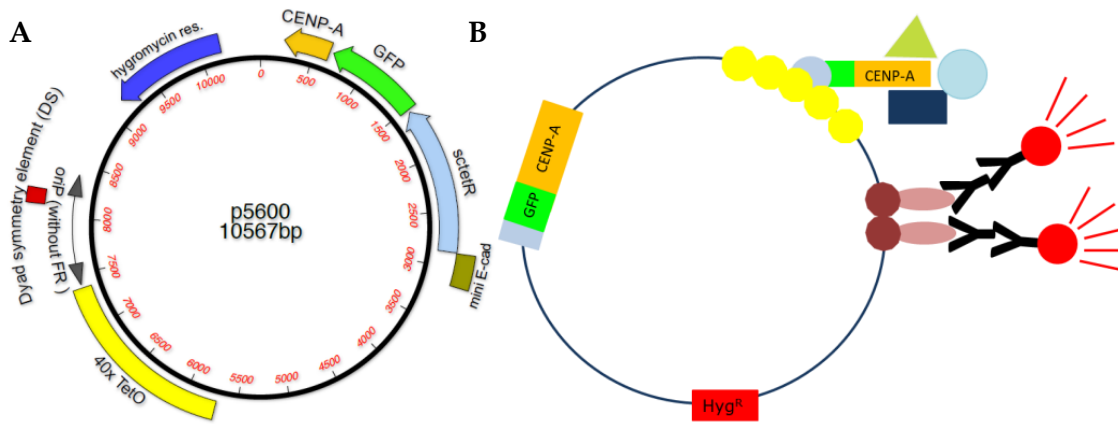
B) ChIP of the tetR fusion protein reveals binding at tetO sites whereas the plasmid region around 5 kb away from tetO is not co-precipitating with scetR:CENP-A. There is a significant enrichment of scetR:CENP-A at tetO compared to the reference region. (n=5, Mean+SD, p-value: unpaired t test with Welch's correction, two-tailed)

scetR:CENP-A was specifically bound to the tetO sequence, comparing the eluted plasmid amount after tetR IP at tetO sites and the reference site (Figure 20). In addition, the tetR pull down with the newly generated  $\alpha$ -tetR antibody is specific, since the control IP with an equivalent IgG subtype only shows background level with the tetO and reference primer pair.

Since I aimed to demonstrate kinetochore protein recruitment by immune fluorescence in the following, I also needed to investigate scetR:CENP-A targeting to plasmids with this method. I know from ChIP experiments, that scetR:CENP-A is site specifically targeted to the tetO array on the plasmids. Therefore, I assume that scetR:GFP:CENP-A co-localizes with the plasmids if they show an overlapping signal with confocal microscopy resolution.

#### 4.2.2 sc tetR:GFP:CENP-A co-localizes with plasmids

In order to illustrate sc tetR:GFP:CENP-A targeting to pCON<sup>CENP-A</sup> in an immune fluorescence approach, I needed to visualize the plasmids. Therefore, the plasmid DNA was represented by the immune fluorescent staining of EBNA1, which interacts with DS, the replication element of the plasmids (Figure 21 B).

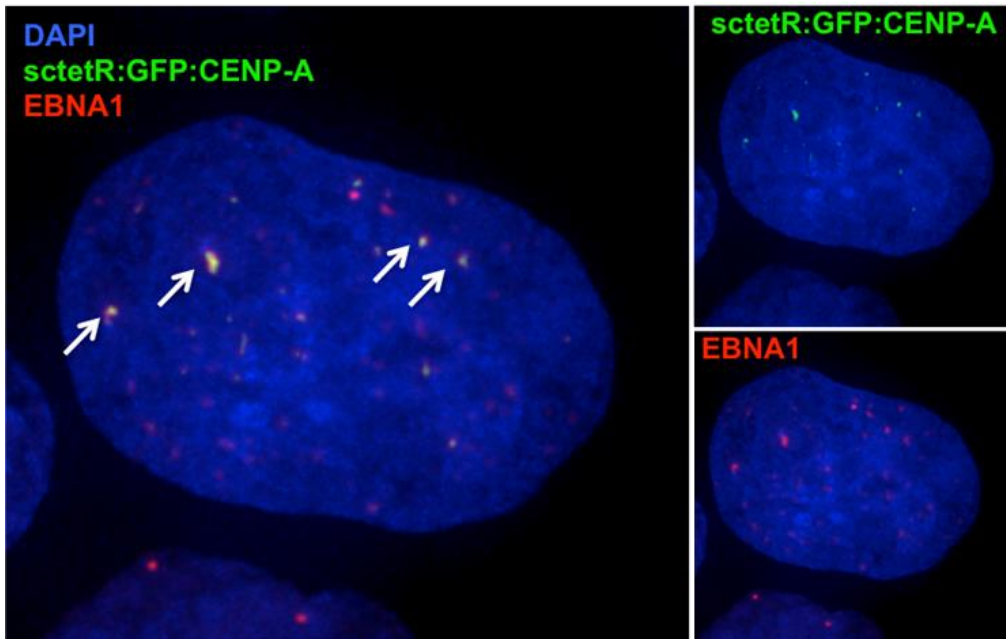


**Figure 21: Illustration of EBNA1 immune fluorescence staining**

A) pCON<sup>CENP-A</sup> reporter plasmid encoding sc tetR:GFP:CENP-A

B) Representation of reporter plasmid with immune fluorescence staining of EBNA1. EBNA1 binds to the DS element of pCON<sup>CENP-A</sup> and is recognized by an EBNA1 specific antibody. This EBNA1 antibody is recognized in turn by fluorescently labelled secondary antibody that recognizes the rat IgG subtype of the EBNA1 specific antibody.

The co-localization of sc tetR:GFP:CENP-A signals and EBNA1 fluorescence staining reveals a close proximity of sc tetR:GFP:CENP-A with the plasmids, which suggests the binding of sc tetR:GFP:CENP-A to the plasmids. In collaboration with Evelyne Barrey, a PhD student in Patrick Heun's group, HEK293EBNA1<sup>+</sup> cells were transfected with pCON<sup>CENP-A</sup> plasmids expressing the sc tetR:GFP:CENP-A fusion protein (Figure 21 A). Four days post transfection, the cells were fixed and stained for the EBNA1 protein using the EBNA1 specific monoclonal 1H4 antibody and a fluorescently labeled secondary antibody raised against the rat IgG subtype of the primary  $\alpha$ -EBNA1 antibody. Co-staining of the sc tetR:GFP:CENP-A with EBNA1 was approved by confocal laser scanning microscopy.



**Figure 22: Immune fluorescence of EBNA1 and sctetR:GFP:CENP-A**  
Co-localization of EBNA1 (representing plasmids) and sctetR:GFP:CENP-A is marked with white arrows. Almost 95 % of EBNA1 signals overlap with sctetR:GFP:CENP-A

Images from Evelyne Barrey (MPI Immunology and Epigenetics, Freiburg)

These representative images suggest the binding of sctetR:GFP:CENP-A to the plasmids (Figure 22, white arrows). Quantification of overlapping signals revealed that almost 95 % of the EBNA1 signals overlap with sctetR:GFP:CENP-A spots (see Figure 29 B and Figure 30 B).

The site-specific targeting of CENP-A to the tetO is the major prerequisite for the establishment of centromere inheritance. In this chapter, I presented that CENP-A is efficiently targeted to pCON<sup>CENP-A</sup> by ChIP and immune fluorescence. The next hypothesis was that the formation of a neo-centromere on plasmids leads to an active segregation mechanism during mitosis. In order to verify this hypothesis, I visualized the plasmids during one cell cycle by live cell imaging.

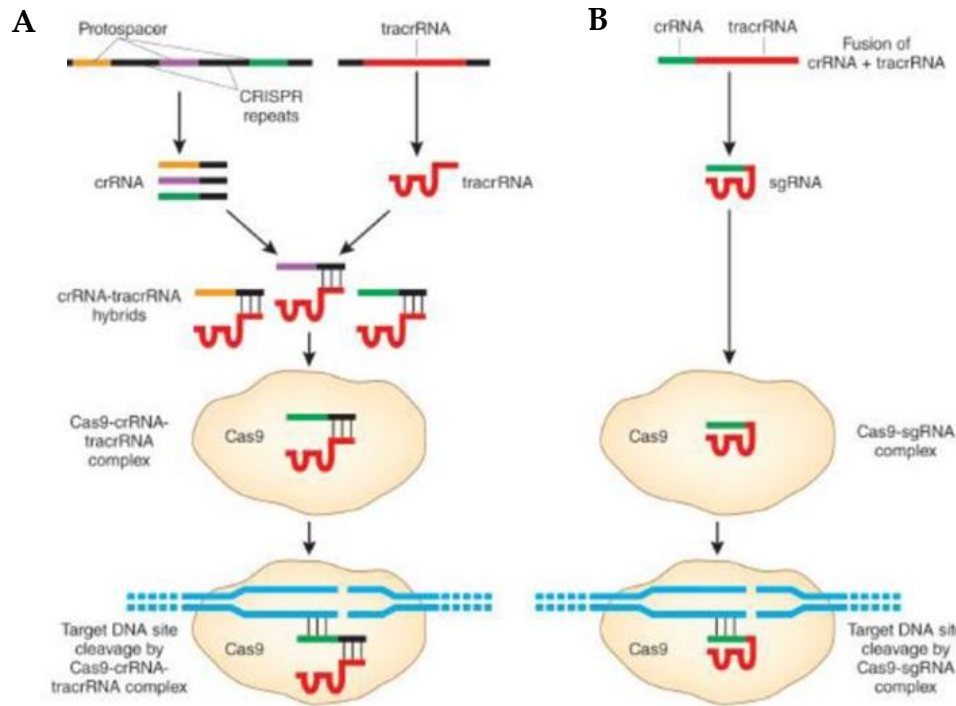


### 4.3 Dynamics of pCON<sup>CENP-A</sup> during cell cycle

For the visualization of plasmid DNA in a live cell imaging approach, the staining of EBNA1 is not suitable. In fixed cells, during the immune fluorescence, EBNA1 signals were amplified by sequential binding of EBNA1 antibodies and even more secondary antibodies that were fluorescently labeled. In live cell imaging, EBNA1 signals are not amplified by sequential antibody staining. In addition, the DS only contains four EBNA1 binding sites and thus the signals resulting from a fluorescently labeled EBNA1 protein are too weak. For that reason, I developed a targeting system to specifically visualize plasmid DNA in live cell settings. Suitable systems are artificial targeting systems, like the prokaryotic LacI-LacO targeting. During this project, I decided to generate an innovative CRISPR/Cas9-dependent visualization method because the CRISPR/Cas9 was demonstrated to visualize specific genomic loci in living cells (Chen et al., 2013).

#### 4.3.1 CRISPR/Cas9 targeting system

The CRISPR (clustered regularly interspaced short palindromic repeats)/Cas9 (CRISPR-associated) system, utilized during this project, is derived from streptococcal defense mechanisms against their pathogens. If foreign DNA enters the cell, it is recognized by the Cas9-crRNA-tracrRNA complex and cleaved by the endonuclease activity of the Cas9 protein. crRNA (CRISPR RNA) is expressed from the CRISPR repeats, containing protospacer sequences. This crRNA interacts with tracrRNA (trans-activating crRNA), which builds a secondary RNA structure that binds to Cas9. After unwinding the foreign DNA, the complementary crRNA binds it and the Cas9 protein cleaves the DNA. (Figure 23 A)



**Figure 23: Naturally occurring and engineered CRISPR-Cas9 systems**

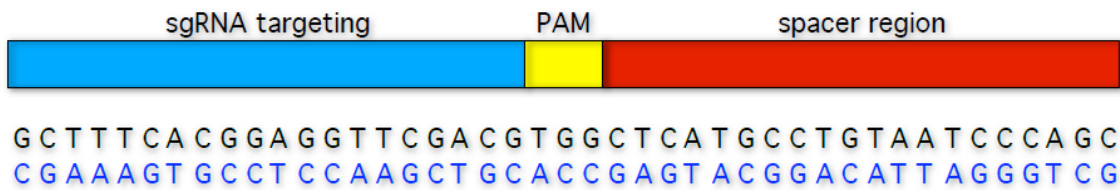
A) Natural CRISPR-Cas9 pathway: foreign DNA sequences are incorporated into CRISPR arrays, which then produce crRNAs including regions that are complementary to the foreign DNA sequence (protospacer). These crRNAs hybridize to tracrRNAs (also encoded in the CRISPR system). This RNA complex associates with the Cas9 nuclease. crRNA/tracrRNA/Cas9 complexes recognize and cleave foreign DNA also complementary to protospacer sequences.

B) Engineered CRISPR-Cas9 system: fusion between a crRNA and part of the tracrRNA sequence (= sgRNA). This sgRNA builds a complex with Cas9 to mediate cleavage of target DNA sites that are complementary to the first (5') 20 nt of the gRNA and that lie next to a PAM sequence.

(Sander and Joung, 2014)

This naturally occurring CRISPR/Cas9 system was engineered to target specific sequences for genome editing or protein targeting. Nuclease active or nuclease deficient Cas9 proteins are used and in both cases, the sgRNA (single guide RNA) is utilized in the same way. The crRNA and the tracrRNA were fused together and thus, the complementary sequence to DNA is directly bound to the RNA forming secondary structure for Cas9 interaction (Figure 23 B). Stable binding of the sgRNA-Cas9 protein complex to DNA occurs if the first 20 nucleotides of sgRNA are complementary to DNA and if the DNA harbors the protospacer adjacent motif (PAM) (Jinek et al., 2012).

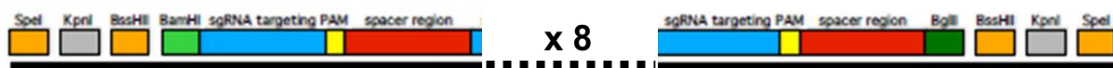
The system, I developed for this project, consisted of a nuclease deficient Cas9 protein fused to three mCherry proteins and a sgRNA that did not have complementary sequences in the human genome. The sgRNA was expressed from a lentiviral vector (clone A66407) and cloned from the library mouse GeCKOv2 library A\_2 by Kai Höfig. I generated a sequence consisting of a complementary sequence to the sgRNA, a PAM and a 20 nucleotide spacer region (Figure 24).



**Figure 24: sgRNA targeting sequence and spacer monomer**

sgRNA complementary sequence for targeting Cas9 fusion proteins for live cell imaging and plasmid purification. For multimerization of this sequence a spacer region of 20 nucleotides was included to avoid interference of neighboring Cas9 proteins.

Spacer region was inserted to reduce interference of neighboring Cas9 proteins that bind to multimers of the sgRNA targeting sequence, since Cas9:3xmCherry is a huge protein. The PAM and the spacer region were chosen from Chen et al. (Chen et al., 2013), supplemental information. A decamer of the sgRNA-PAM-spacer sequence was designed with MacVector and ordered as GeneScript plasmid. For further multimerization to 20, 30 and 40 repeats and cloning into reporter plasmids, the decameric sequence was equipped with specific restrictions sites (Figure 25).



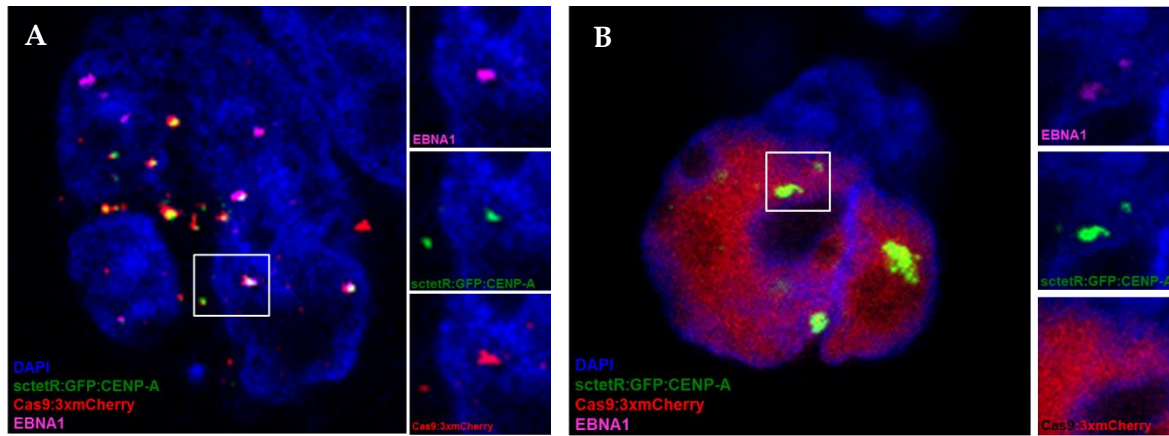
**Figure 25: sgRNA targeting sequence with restriction sites**

Decamer of sgRNA targeting, PAM and spacer region was equipped with specific restriction sites for further multimerization (BamHI, BglII) and final cloning into reporter plasmids (SpeI, KpnI, BssHII)

The live cell imaging application of Cas9:3xmCherry targeting and the dynamics of pCON<sup>CENP-A</sup> during the cell cycle are described in the following.

#### 4.3.2 Cas9:3xmCherry is suitable to visualize pCON<sup>CENP-A</sup> plasmids

In order to use Cas9:3xmCherry targeting in live cell imaging microscopy, I first demonstrated specific targeting of Cas9:3xmCherry proteins to the binding sites on the plasmids. Therefore, I co-transfected HEK293EBNA1<sup>+</sup> cells with the Cas9:3xmCherry and the sgRNA expression plasmids together with pCON<sup>CENP-A</sup> with and without the specific sgRNA complementary sequence. After four days, I fixed those cells and performed EBNA1 immune fluorescent staining to represent the plasmids, like in chapter 4.2.2.



**Figure 26: Cas9:3xmCherry is specifically targeted to plasmids in IF**

A) sctetR:GFP:CENP-A and Cas9:3xmCherry co-localize with plasmid signals

B) No co-localization of Cas9:3xmCherry with plasmids and sctetR:GFP:CENP-A, if plasmid does not contain Cas9 targeting sites. Cas9:3xmCherry is distributed in the whole nucleus.

pCON<sup>CENP-A</sup> plasmids containing Cas9 specific targeting site, represented by EBNA1 signals, overlapped with the Cas9:3xmCherry fluorescence. In addition, many signals arising from sctetR:GFP:CENP-A co-localized with Cas9:3xmCherry if the plasmids harbored complementary sgRNA sequence repeats (Figure 26 A). In contrast, if pCON<sup>CENP-A</sup> does not contain the targeting sites, the Cas9 fusion protein was distributed equally in the whole nucleus (Figure 26 B).

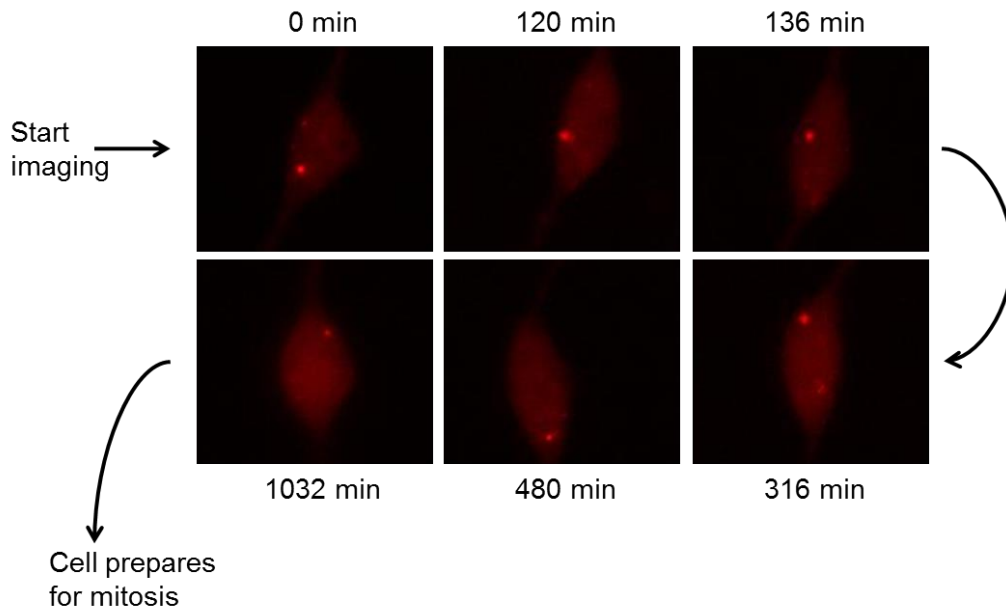
Only the presence of specific targeting sites for the sgRNA leads to the specific localization of the Cas9:3xmCherry fusion protein to plasmids. This makes the Cas9

artificial targeting system a suitable tool to investigate plasmid dynamics in living cells by live cell imaging.

### 4.3.3 pCON<sup>CENP-A</sup> localization in living cells

To understand the pCON<sup>CENP-A</sup> segregation mechanism in mitosis and plasmid localization during interphase, live cell imaging was performed. Imaging living and proliferating cells has to deal with some critical aspects. Excitation with energetic wavelengths and high intensities leads to increased cell death. In addition, oxygen and CO<sub>2</sub> concentration and humidity need to be optimized for imaging conditions of proliferating cells. For that reason I decided to collaborate with Andreas Thomae and Steffen Dietzel from the Biomedical Center Munich, since they were experienced and equipped for live cell imaging. With LED driven fluorescence excitation, I performed live cell imaging in a suitable time resolution. By keeping excitation intensities low, cells survived and proliferated under imaging conditions, even though imaging up to three different fluorescent colors.

For live cell imaging I transfected a HEK293EBNA1<sup>+</sup> cell line stable expressing Cas9:3xmCherry with the pCON<sup>CENP-A</sup> reporter plasmid, already used in the fixed cells. After four days establishment time, the cells were transferred onto an imaging slide and imaged at a fluorescence live cell microscope in cooperation with Andreas Thomae. Every four minutes an image stack was taken from bright field illumination, the scetR:GFP:CENP-A and the Cas9:3xmCherry channel for 24 hours in total. Signals of scetR:GFP:CENP-A were too weak to follow over one cell cycle and only the dynamics of Cas9:3xmCherry were analyzed.

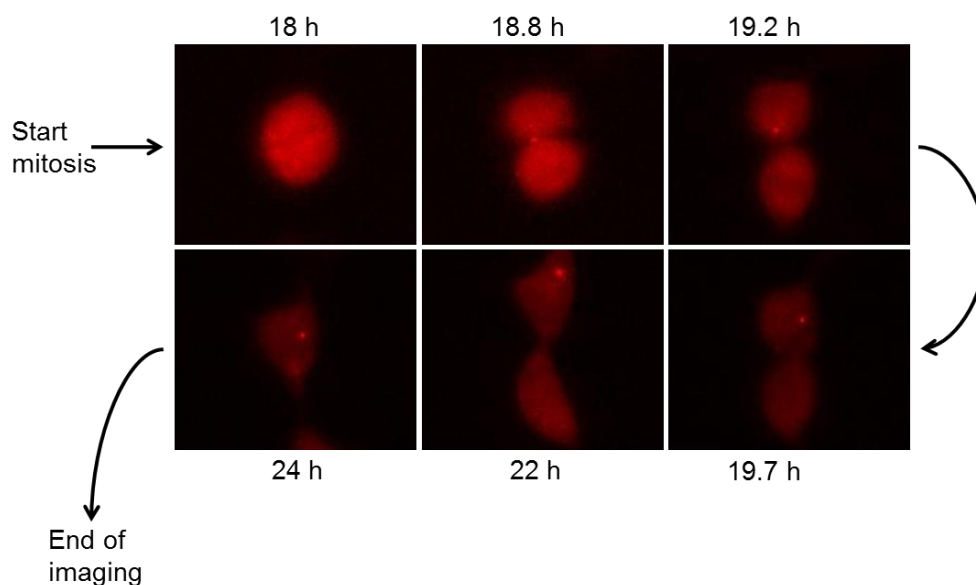


**Figure 27: Highly dynamic localization of pCON<sup>CENP-A</sup> in interphase**  
Plasmids are stained by Cas9:3xmCherry. Images were taken every 4 min for 24 hours. Different time points during interphase are selected here

During interphase the plasmid signals were highly dynamic. The bright Cas9:3xmCherry signal representing plasmids that was followed in Figure 27, moved from one end of the nucleus to another and back. Consequently, the plasmids are not immobilized within any specific environment, but rather move through the whole nucleus.

Even if the cell prepared for mitosis, the plasmid localization was not static.

During mitosis the plasmid signal localized in between the future daughter cells during anaphase (Figure 28, 18.8 h). After cell division the plasmid signal, represented by Cas9:3xmCherry, was observed in one of both daughter cells. This means that in this example shown here, it seems that plasmids are segregated asymmetrically during mitosis.



**Figure 28: plasmid segregation in mitosis is asymmetric**

Same imaging time series as in Figure 28. Cell imaged during mitosis. In this cell plasmid aggregate is asymmetrically distributed to just one daughter cell.

In several other experiments I aimed to further characterize segregation of plasmids having an artificial centromere, but unfortunately I couldn't manage to visualize living cells expressing Cas9:3xmCherry anymore. In additional experiments I stained living cells with the live-cell tubulin dye to demonstrate attachment of microtubules to plasmids. Because of technical issues and limited time, I was not able to reveal the exact segregation mechanism by live cell imaging during this project.

In summary, the Cas9-dependent targeting system is specifically targeted to the plasmids, demonstrated by the co-localization to pCON<sup>CENP-A</sup> plasmids in fixed cells. Therefore, it is a suitable tool for live cell investigations of plasmid dynamics. Nevertheless, the segregation mechanism of the reporter plasmids was not clarified with this method.

However, we supposed active plasmid segregation because of the establishment of a neo-centromere on the pCON<sup>CENP-A</sup> plasmids. To verify this, I examined recruitment of inner and outer kinetochore components since the kinetochore assembly is the prerequisite of microtubule mediated active segregation. If the kinetochore is assembled on plasmids it is likely that these are segregated actively during mitosis.

### 4.4 Recruitment of kinetochore proteins

We observed that centromere establishment on plasmids after CENP-A targeting led to maintenance of these plasmids as independent genetic entities by centromere regulated segregation. An active plasmid segregation mechanism implies the recruitment of kinetochore proteins to the plasmids, as they mediate the interaction of centromeric chromatin with the microtubules.

I investigated the CENP-A dependent recruitment of kinetochore proteins to the artificial centromere on the plasmid by immune fluorescence. In the following I delineate the inner kinetochore protein CENP-C and the outer kinetochore protein Ndc80 recruitment to neo-centromeres after CENP-A targeting.

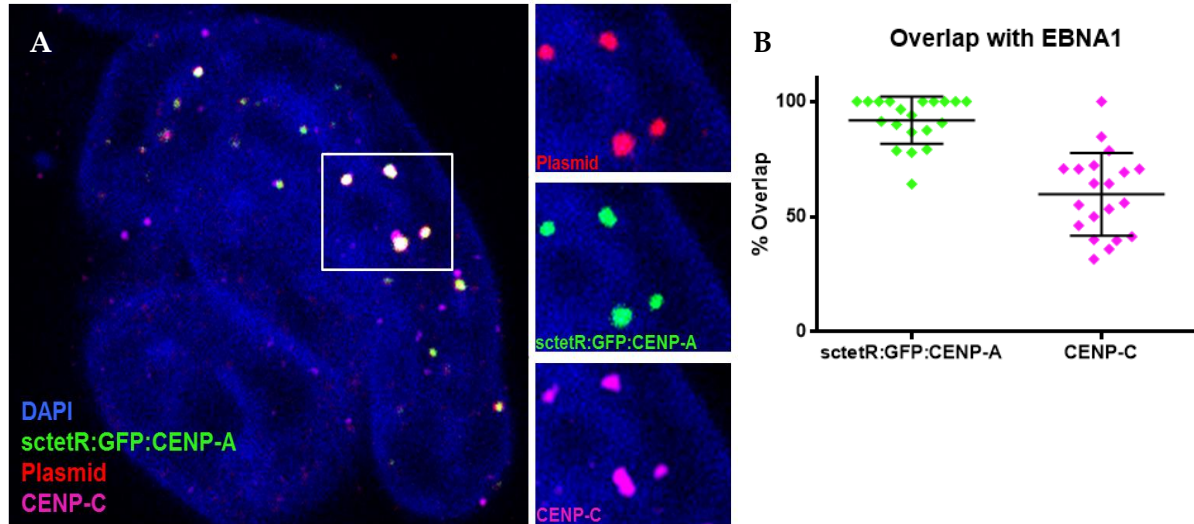
#### 4.4.1 Recruitment of inner kinetochore proteins (CENP-C)

In an immune fluorescence approach I studied if the recruitment of the inner kinetochore protein CENP-C to pCON<sup>CENP-A</sup> is dependent on CENP-A targeting. CENP-C has some important functions at endogenous centromeres. It is involved in the assembly, formation and maintenance of kinetochores (Tomkiel et al., 1994). CENP-C is directly recruited and bound by CENP-A cell cycle independently and it reshapes and stabilizes CENP-A nucleosomes (Falk et al., 2015). To fulfil these functions on pCON<sup>CENP-A</sup> as well, CENP-C has to be recruited to the plasmids after CENP-A targeting.

To verify the recruitment of CENP-C to reporter plasmids immune fluorescence staining of EBNA1 and CENP-C were performed. HEK293EBNA1<sup>+</sup> cells were transfected with the pCON<sup>CENP-A</sup> plasmid expressing scetR:GFP:CENP-A, as described in 4.2.2. Four days after transfection, cells were fixed and stained for EBNA1 and CENP-C with specific antibodies. Confocal microscopy images were taken of 20 cells. Interphase as well as mitotic cells were used to quantify the co-



localization of CENP-C with the plasmids, since CENP-C is constitutively bound to CENP-A.



**Figure 29: Recruitment of the inner kinetochore protein CENP-C**

A) A representative image is shown for the recruitment of CENP-C to plasmids with targeted sctetR:GFP:CENP-A. Enlarged spots demonstrate overlapping CENP-C signals with plasmids, represented by EBNA1 staining.

B) Quantification of CENP-C/EBNA1 co-localization on plasmids. First EBNA1 spots were marked then CENP-C and sctetR:GFP:CENP-A overlap on these spots was analyzed. 20 cells were used for quantification and sctetR:GFP:CENP-A or CENP-C signals were calculated relative to the EBNA1 signals (Mean+SD).

sctetR:GFP:CENP-A targeting and formation of centromeres on pCON<sup>CENP-A</sup> led to recruitment of CENP-C, since sctetR:GFP:CENP-A, plasmids, represented by EBNA1, and CENP-C fluorescence depicts overlapping signals in immune fluorescence microscopy (Figure 29 A).

Artificial targeting of sctetR:GFP:CENP-A and recruitment of CENP-C to the plasmids was quantified from 20 cells (Figure 29 B). In order to quantify the targeting of CENP-A and the co-localization of CENP-C with the artificial centromeres, EBNA1 signals were determined first. Second, the signal overlap of sctetR:CENP-A with the EBNA1 signals was analyzed. As represented in Figure 29 B (green representation) on average 95 % of EBNA1 signals co-localize with sctetR:GFP:CENP-A. These 95 % of overlapping signals, were then scanned for CENP-C co-staining. It turned out that 55 % of plasmids recruit CENP-C (Figure 29 B, magenta representation). No co-

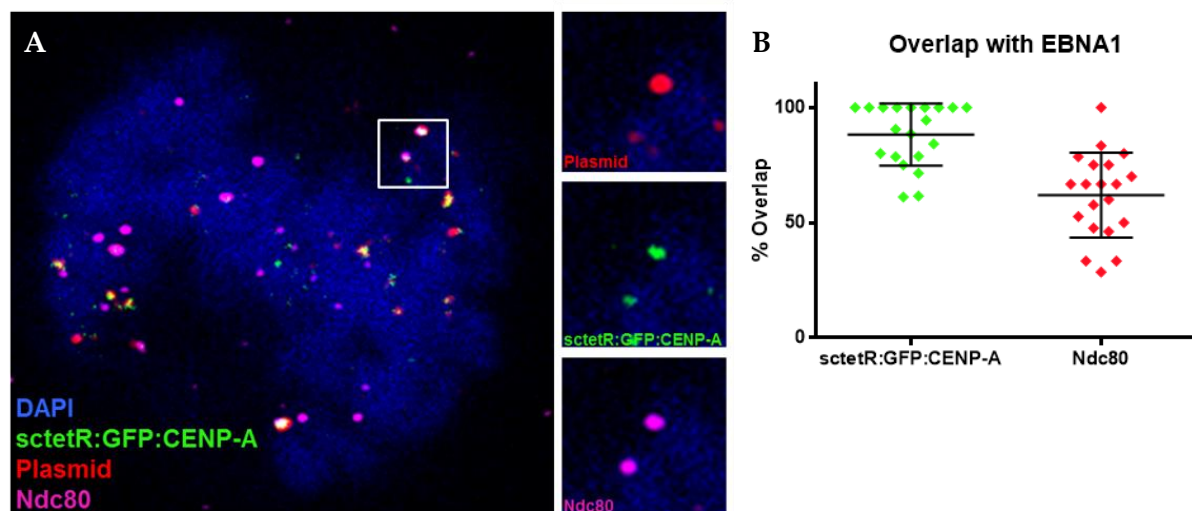
staining of CENP-C and EBNA1 was observed when sctetR:GFP:CENP-A signal was not present.

By immune fluorescence microscopy I demonstrated a recruitment of CENP-C to pCON<sup>CENP-A</sup> that is dependent on CENP-A targeting. CENP-C is one major factor for the assembly of the outer kinetochore complex during mitosis. Since, this protein was present at the neo-centromeres on plasmids, we suggested that also outer kinetochore components assemble during mitosis.

### 4.4.2 Recruitment of outer kinetochore proteins (Ndc80)

As second example to confirm the hypothesis that centromere establishment on plasmids after CENP-A targeting leads to active plasmid segregation by kinetochore assembly on pCON<sup>CENP-A</sup>, I studied the co-staining of EBNA1 with the outer kinetochore component Ndc80 by immune fluorescence. Since, outer kinetochore proteins and microtubules are recruited to centromeres during mitosis, I had to analyze mitotic cells. HEK293EBNA1<sup>+</sup> cells were transfected with the reporter plasmid and four days after transfection, cells were synchronized by thymidine block. Cells were released from block and 8 hours after release I fixed them during mitosis. After immune fluorescent staining of Ndc80 and EBNA1, the signal overlap of Ndc80 and EBNA1 of 20 cells was quantified as for the CENP-C recruitment.

During mitosis the outer kinetochore protein Ndc80 was recruited to the plasmids (Figure 30 A). The co-localization of sctetR:GFP:CENP-A with EBNA1 in mitotic cells was similar to that in interphase cells. Around 95 % of plasmid signals overlap with sctetR:GFP:CENP-A fluorescence during mitosis (Figure 30 B, sctetR:GFP:CENP-A). On average 60 % of these plasmids recruit the outer kinetochore protein Ndc80 (Figure 30 B, Ndc80).



**Figure 30: Recruitment of the outer kinetochore protein Ndc80**

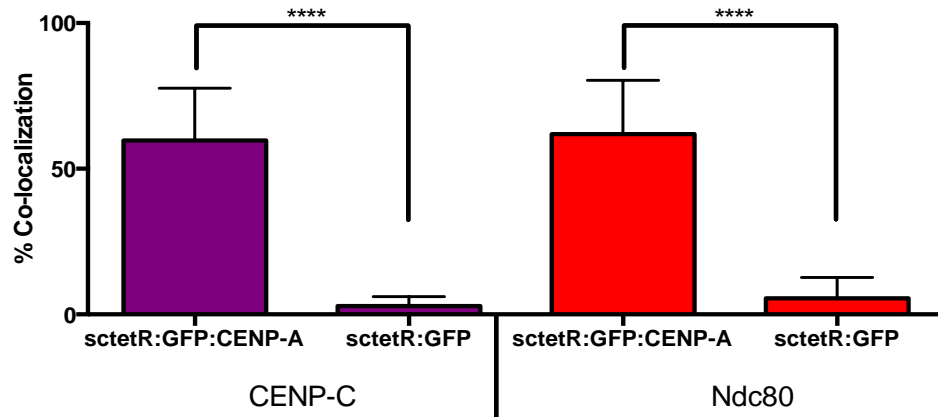
A) A representative image of the recruitment of Ndc80 to reporter plasmids with targeted sctetR:GFP:CENP-A during mitosis is shown. Enlarged spots demonstrate overlapping signals of Ndc80 and EBNA1 fluorescent staining.

B) Quantification of Ndc80/EBNA1 co-localization on plasmids. First EBNA1 spots were marked then Ndc80 and sctetR:GFP:CENP-A overlap on these spots was analyzed. 20 cells were used for quantification and sctetR:GFP:CENP-A or Ndc80 signals were calculated relative to the EBNA1 signals (Mean+SD).

In immune fluorescence experiments, I demonstrated the recruitment of inner and outer kinetochore components after artificial targeting of CENP-A. Since I observed that EBNA1 signals, not overlapping with sctetR:GFP:CENP-A, also do not recruit kinetochore proteins, I assume that recruitment of CENP-C and Ndc80 to plasmids is dependent on CENP-A targeting.

#### 4.4.3 Recruitment of kinetochore proteins only if CENP-A is targeted

In order to demonstrate that the recruitment of kinetochore proteins to pCON<sup>CENP-A</sup> is dependent CENP-A targeting, a control experiment was performed. I transfected HEK293EBNA1<sup>+</sup> cells with pCON<sup>CENP-A</sup> and a tetO-DS plasmid expressing the sctetR:GFP fusion protein. I quantified the overlap of plasmid signals, represented by the EBNA1-DS interaction, with sctetR:GFP and CENP-C or Ndc80 of 20 cells.



**Figure 31: Quantification of kinetochore protein co-localization after targeting sctetR:GFP:CENP-A or sctetR:GFP**

Comparison of quantification of CENP-C and Ndc80 co-localization with plasmids targeted by either sctetR:GFP:CENP-A or sctetR:GFP. Quantification was carried out as described in Figure 29. (n=20 cells, Mean+SD,  $p < 0.0001$ , p-value: unpaired t test with Welch's correction, two-tailed)

The recruitment of CENP-C and Ndc80 was dependent on the targeting of CENP-A, because artificial targeting sctetR:GFP to tetO-DS reporter plasmids did not result in co-staining of EBNA1, the sctetR fusion protein and kinetochore proteins. In contrast, if sctetR:GFP:CENP-A was targeted, a highly significant increase in signal overlap between EBNA1 and kinetochore proteins was observed.

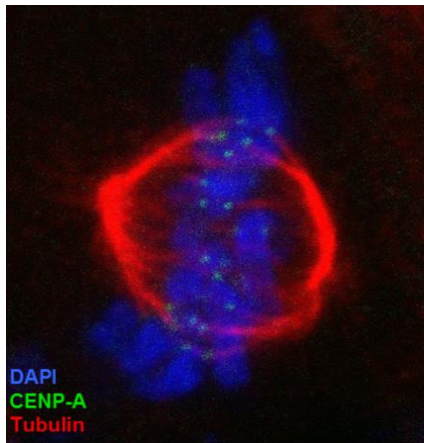
I conclude from these results that artificial CENP-A targeting leads to the establishment of a functional centromere identity on  $pCON^{CENP-A}$ . Kinetochore components, represented by CENP-C and Ndc80, are recruited to neo-centromeres in a CENP-A dependent manner. During mitosis the outer kinetochore mediates active plasmid segregation and this in turn leads to long-term plasmid maintenance.

To gain deeper understanding in the mechanisms that lead to establishment of centromere inheritance on  $pCON^{CENP-A}$ , I investigated the capacity of sctetR:CENP-A fusion proteins to form centromeric nucleosomes.

## 4.5 sctetR:CENP-A forms centromeric nucleosomes

The sctetR:CENP-A fusion protein is targeted to the tetO array of pCON<sup>CENP-A</sup> plasmids by the interaction of the tet transactivator (tetR) and its target site. This interaction was already confirmed in previously described experiments (chapter 4.2). In addition, the recruitment of kinetochore components to the artificial centromere on plasmids by targeting CENP-A was demonstrated. This suggests that the CENP-A fusion protein is indeed forming a kinetochore complex at an artificial centromere. The following experiments aim to verify this hypothesis, by analyzing if the fusion protein is incorporated into nucleosomes.

### 4.5.1 sctetR:GFP:CENP-A is present at endogenous centromeres



**Figure 32: Immune fluorescence of tubulin**

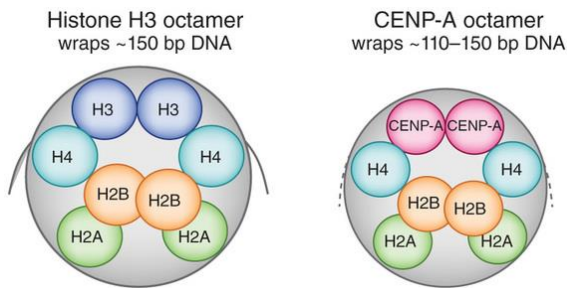
Binding of microtubules at endogenous centromeres where also sctetR:GFP:CENP-A is bound.

To address the question, if sctetR:CENP-A is recruited to endogenous centromeres, a sctetR:GFP:CENP-A expression plasmid was transfected into HEK293EBNA1<sup>+</sup> cells. Four days after transfection, cells were fixed and immune fluorescently stained against tubulin proteins. The attachment of microtubules to sctetR:GFP:CENP-A containing centromeres was detected by confocal microscopy. With this approach I demonstrated that the sctetR:GFP:CENP-A fusion protein was recruited to endogenous centromeres (Figure 32). During mitosis microtubules attached at endogenous centromeres and sctetR:GFP:CENP-A signals were present at microtubule ends.

The actual incorporation of sctetR:CENP-A fusion proteins into centromeric nucleosomes was not shown by immune fluorescence. In order to clarify the ability

of sctetR:CENP-A to incorporate into nucleosomes, a co-immuno precipitation (co-IP) was performed by demonstrating interaction of sctetR:CENP-A fusion proteins with canonical histones.

#### 4.5.2 sctetR:CENP-A nucleosomes



**Figure 33: schematic representation of H3 and CENP-A nucleosomes**

Canonical histone only contains H3 histones, whereas CENP-A nucleosomes only contain CENP-A histones

Adopted from Dunleavy & Karpen, *Nature structural and molecular biology* (2013) (Dunleavy et al., 2013)

As already described in detail in chapter 1.5.2, the centromeric CENP-A nucleosome varies in structure and composition from canonical nucleosomes. The CENP-A histone H3 variant dimerizes with histone H4 to a (CENP-A:H4)<sub>2</sub> heterodimer. These are incorporated, together with two H2A and H2B histones into centromeric DNA. This complex builds an octameric nucleosome within the centromeres (Figure 33).

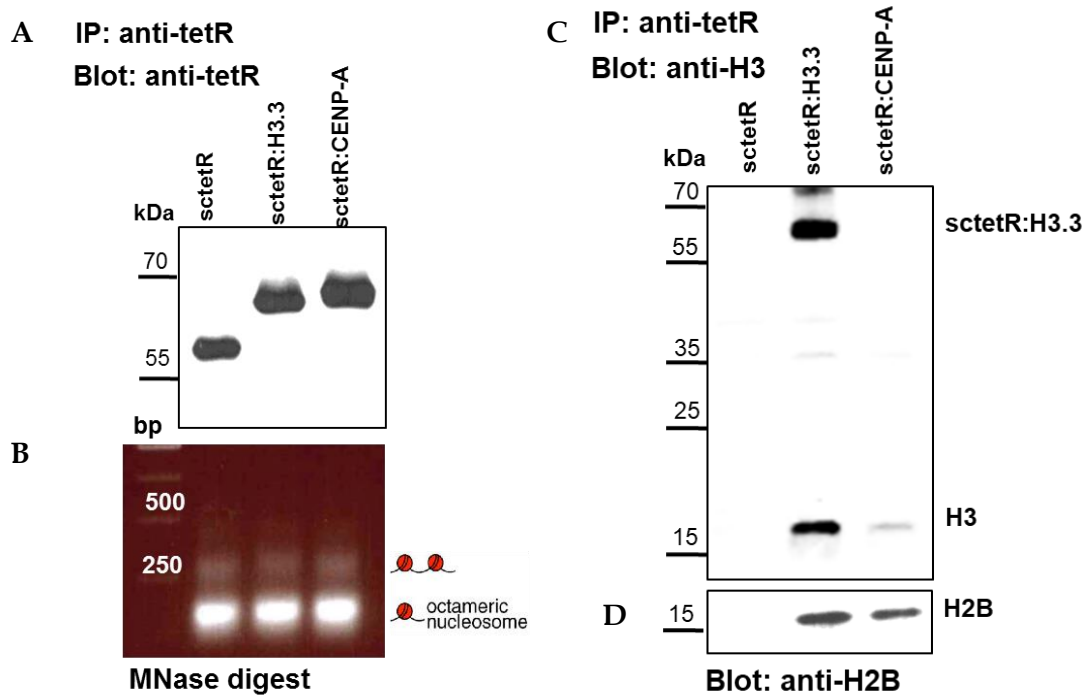
Other than canonical nucleosomes, centromeric nucleosomes contain two CENP-A histone H3 variants instead of canonical H3 (Dunleavy et al., 2013).

To analyze whether sctetR:CENP-A is incorporated into centromeric nucleosomes an immuno precipitation (IP) of the tetR fusion protein was performed and the co-precipitation of canonical histones was investigated.

Mono-nucleosomes (Figure 34 B) were extracted from three different cell lines (HEK293EBNA<sup>+</sup> + sctetR, sctetR:H3.3 or sctetR:CENP-A). sctetR fusion proteins were precipitated with the tetR specific 31B3 antibody covalently coupled to protein G beads. After stringent washing steps, precipitates were examined for co-precipitation of canonical H3 and H2B.



If sctetR:CENP-A is incorporated correctly into CENP-A nucleosomes, a co-precipitation of H2A, H2B and H4 but not H3 is expected. In contrast, in cells expressing sctetR without fusion, no histone co-precipitates after a tetR IP of mono-nucleosomes. In case of sctetR:H3.3 cell line, histones H2A, H2B, H4, endogenous H3 but not CENP-A are expected to co-precipitate.



**Figure 34: tetR IP and co-IP of endogenous histones**

- A) IP analysis of tetR IP from different cell lines sctetR, sctetR:H3.3 and sctetR:CENP-A  
 B) MNase digest, in all cell lines DNA is digested to mono-nucleosomes  
 C) Western Blot of eluted fractions against histone H3. Only in sctetR:H3.3 cell line H3 is co-precipitating. (n=1)  
 D) Western Blot of eluted fractions against H2B. H2B co-precipitates with sctetR:H3.3 and sctetR:CENP-A. (n=1)

I obtained a comparable sctetR fusion protein amounts after the tetR-IP of sctetR, sctetR:H3.3 and sctetR:CENP-A from the respective cell lines (Figure 34 A). For the IP of sctetR alone no co-precipitation of endogenous histone proteins was observed (Figure 34 C, D; sctetR). In contrast, in case of sctetR fused to H3.3, endogenous H3 and H2B were co-precipitating from MNase digested extract (Figure 34 C, D; sctetR:H3.3). This indicates that the sctetR:H3.3 fusion protein is incorporated into canonical nucleosomes. With the sctetR:CENP-A fusion protein only H2B, but almost

no H3 was co-precipitated (Figure 34 C, D; sctetR:CENP-A). This result confirms that the sctetR:CENP-A complexes with H2A, H2B and H4 to form centromere-like nucleosomes.

The results described in this chapter demonstrate that the sctetR:CENP-A fusion protein is incorporated into centromeric nucleosomes. The fusion protein interacts with the other histones and forms centromere specific nucleosome particles, revealed by co-precipitation. Furthermore, the sctetR:GFP:CENP-A fusion protein is present at endogenous centromeres, where microtubules are associated during mitosis.

Since the kinetochore is assembled on pCON<sup>CENP-A</sup> as well, and plasmid maintenance is stable over several months, I assume that sctetR:CENP-A is incorporated as centromere-like nucleosome into the plasmids. This incorporation leads to the establishment of an inheritable and self-propagating centromere. To verify that hypothesis, I studied the dependence of plasmid maintenance on constitutive targeting of CENP-A.

### 4.6 Establishment of centromere inheritance

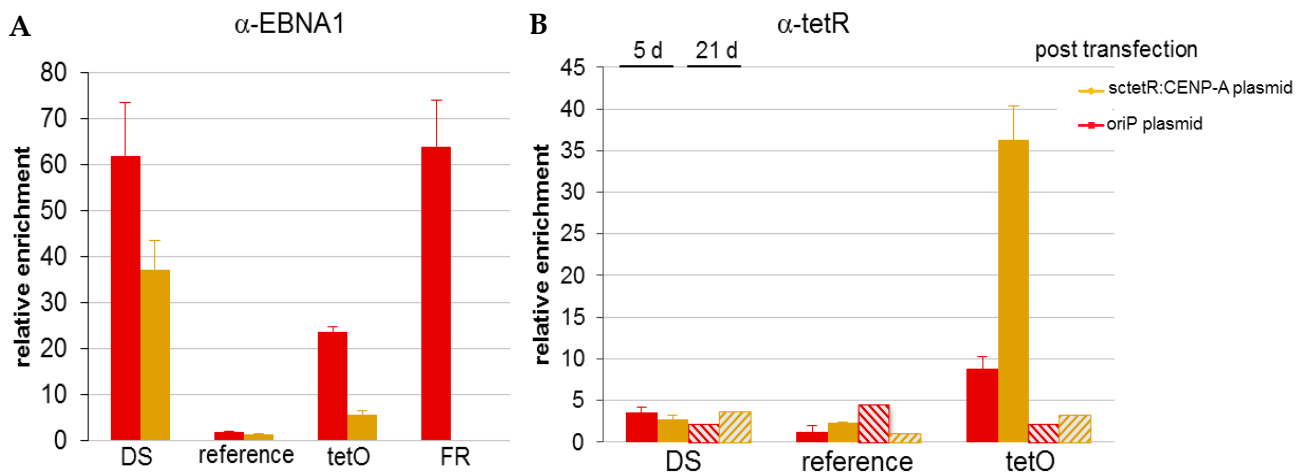
In *Drosophila* cells our group demonstrated, together with our collaboration partner Patrick Heun, that initial CENP-A targeting induces an inheritable centromere on plasmids that becomes independent on CENP-A targeting within three weeks (Mendiburo et al., 2011). Since we know that artificial CENP-A targeting in human cells also leads to the recruitment of kinetochore proteins and therefore to the establishment of a neo-centromere, we studied the establishment of the centromere inheritance in more detail. We aimed to ascertain the dependence of plasmid centromeres on CENP-A targeting, the timeframe of *de novo* centromere establishment on plasmids and its ability to self-propagate.



#### 4.6.1 Targeting of sctetR:CENP-A after 5 days but not after 3 weeks

In order to analyze if plasmid maintenance is dependent on constitutive CENP-A targeting we performed a chromatin immuno precipitation (ChIP) of sctetR:CENP-A at different time points after transfection.

Stefanie Fülöp, a former group member, transfected HEK293EBNA1<sup>+</sup> cells with *oriP* and pCON<sup>CENP-A</sup> and harvested the cells five and 21 days after transfection. EBNA1 binding to DS and FR on *oriP* and to DS on pCON<sup>CENP-A</sup> was determined five days after transfection, the binding of sctetR:CENP-A to tetO was compared five and 21 days after transfection by ChIP.



**Figure 35: ChIP of EBNA1 and sctetR:CENP-A in *oriP* and pCON<sup>CENP-A</sup> plasmid**

A) EBNA1 is bound to DS and FR in the *oriP* plasmid and it is bound to DS in the pCON<sup>CENP-A</sup> plasmid system.

B) sctetR is bound to tetO sites 5 days after transfection. It is not bound to other regions of plasmid. Three weeks after transfection, there is no binding of sctetR:CENP-A to tetO.

Data: Stefanie Fülöp (unpublished)

Analysis of EBNA1 ChIP revealed that the EBNA1 protein was bound to the DS element, present on both reporter plasmids, with high affinity. As expected, the same enrichment of EBNA1 was found at the FR element of *oriP*. Whereas no enrichment of EBNA1 at the tetO and FR primer region on the sctetR:CENP-A plasmid was detected. Since the reference region is several kilo base pairs away from EBNA1

binding sites, no significant enrichment of EBNA1 was observed in this region (Figure 35 A). In addition, tetR ChIP identified background level enrichment at DS, tetO and the reference regions on *oriP* (Figure 35 B, red bars).

sctetR:CENP-A enrichment at tetO and reference regions (DS and reference) was determined five days and three weeks after transfection (for primer positions see Figure 20). Quantification of sctetR:CENP-A binding to tetO five days after transfection revealed an enrichment of sctetR:CENP-A at tetO repeats (Figure 35 B, yellow, filled bars). In contrast, analyzing sctetR:CENP-A binding after three weeks, no sctetR:CENP-A enrichment at the tetO array was detected (Figure 35 B, yellow, dashed bars).

This observation is a first hint, that after initial targeting of sctetR:CENP-A maturation of the centromere identity takes place on the plasmid leading to a sctetR:CENP-A targeting independent maintenance.

The time needed for establishment of centromere identity and the maturation of centromeric chromatin on the artificial plasmid centromere is addressed in the following.

### 4.6.2 Centromere inheritance is established within 7 days

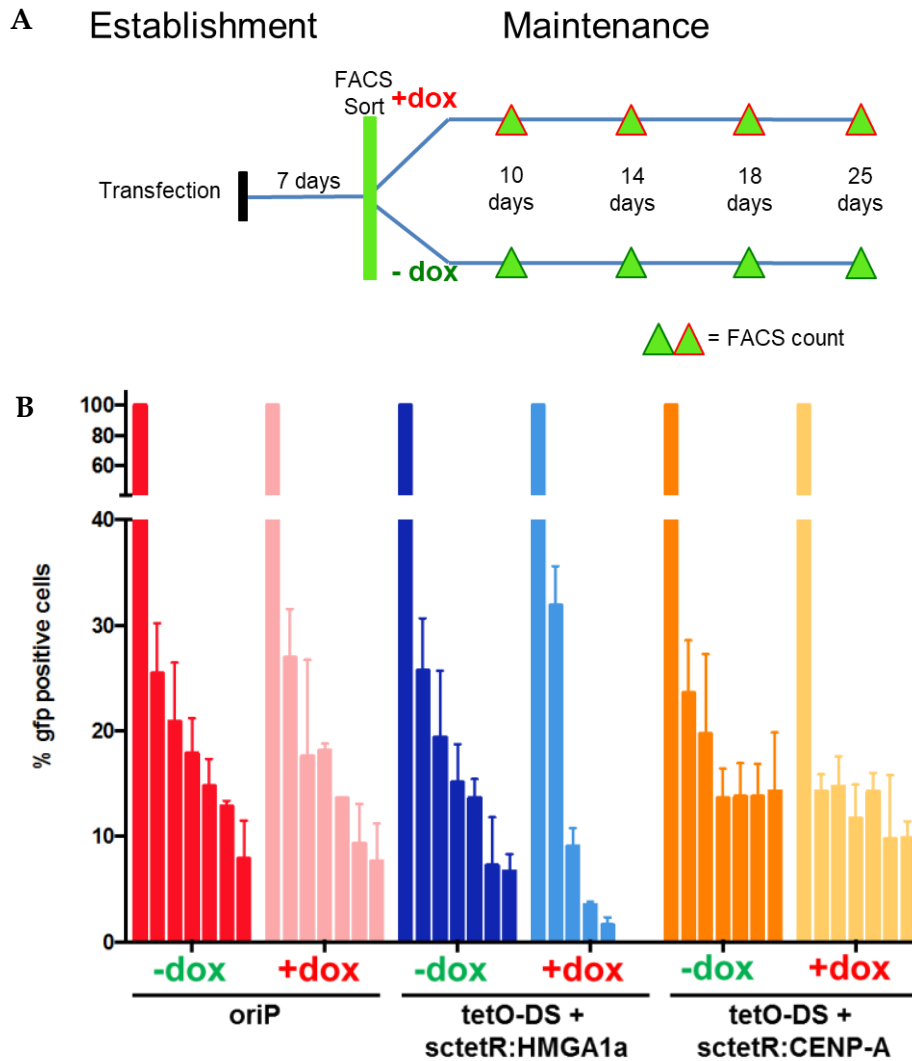
In our previous experiments, we investigated plasmid maintenance after an establishment phase of two to four weeks under selective pressure. Since we observed the recruitment of kinetochore proteins takes place earlier, I performed, in close collaboration with Lara Schneider, plasmid maintenance experiments analyzing the establishment of centromere identity after seven days.

For this purpose, we made use of doxycycline, that inhibits binding of sctetR to the tet operator sequence by changing the conformation of the tet transactivator dimer. We analyzed plasmid maintenance after a shortened establishment time of seven days and after adding doxycycline to the cells. To determine whether plasmid

maintenance of the *sctetR:CENP-A* plasmids is dependent on *sctetR:CENP-A* targeting, we also compared it with two different controls. The *oriP* plasmid maintenance is only dependent on EBNA1 and thus not influenced by doxycycline. In contrast, the maintenance of *sctetR:HMGA1a* plasmids, introduced in chapter 1.8, is strictly dependent on constitutive targeting on *sctetR:HMGA1a* and therefore the plasmid loss is inducible by doxycycline (Pich et al., 2008). This doxycycline inducible system served as control for doxycycline activity. In the “+dox” condition the *sctetR:HMGA1a* plasmids show an increased plasmid loss compared to the “-dox” condition.

HEK293EBNA1<sup>+</sup> cells were transfected with the different reporter plasmids (*oriP*, *sctetR:HMGA1a tetO-DS* and *pCON<sup>CENP-A</sup>*) and incubated for seven days. After this incubation time (= establishment phase) cells were FACS sorted according to *gfp* expression. Only *gfp* positive cells were re-plated and cultivated in presence and absence of 2 µg/ml doxycycline in cell growth medium. Plasmid maintenance was then analyzed by FACS count for *gfp*<sup>+</sup> at different time points after the FACS sort (Figure 36 A).

As expected *oriP* plasmid loss was independent of doxycycline treatment. A doxycycline inducible plasmid loss was observed for the *sctetR:HMGA1a* plasmid system. Plasmid loss rates of *oriP* and *sctetR:HMGA1a tetO-DS* without doxycycline were similar, whereas the inhibition of *sctetR:HMGA1a* binding to tetO by doxycycline led to a significant increase in plasmid loss after seven days establishment time. Within 18 days only low levels of *gfp*<sup>+</sup> cells were detectable. In contrast the *pCON<sup>CENP-A</sup>* plasmids were maintained at a stable level in both tested conditions (“+dox” and “-dox”). About 15 % *gfp*<sup>+</sup> cells were the stable threshold for maintaining *pCON<sup>CENP-A</sup>* plasmids after an initial plasmid loss, independent of doxycycline.



**Figure 36: Targeting independent plasmid maintenance 7 days after establishment**

A) Timeline of experimental setup. 7 days after transfection cells are sorted according to gfp expression and then splitted into condition with and without doxycycline. Plasmid maintenance was analyzed by FACS count every 4 to 6 days.

B) Plasmid maintenance after 7 days establishment time. (n=3, Mean+SD)

We found that initial targeting of CENP-A for seven days is sufficient to establish inheritable centromeres on pCON<sup>CENP-A</sup>. The same experimental setup was then performed with four days establishment phase (see Appendix). Four days of establishment showed the same plasmid maintenance as seven days (compare Figure 36 and Appendix Figure 5). These results raised the question if the establishment of an artificial centromere can be even faster than four days.

### 4.6.3 Establishment of centromere inheritance after 4 days

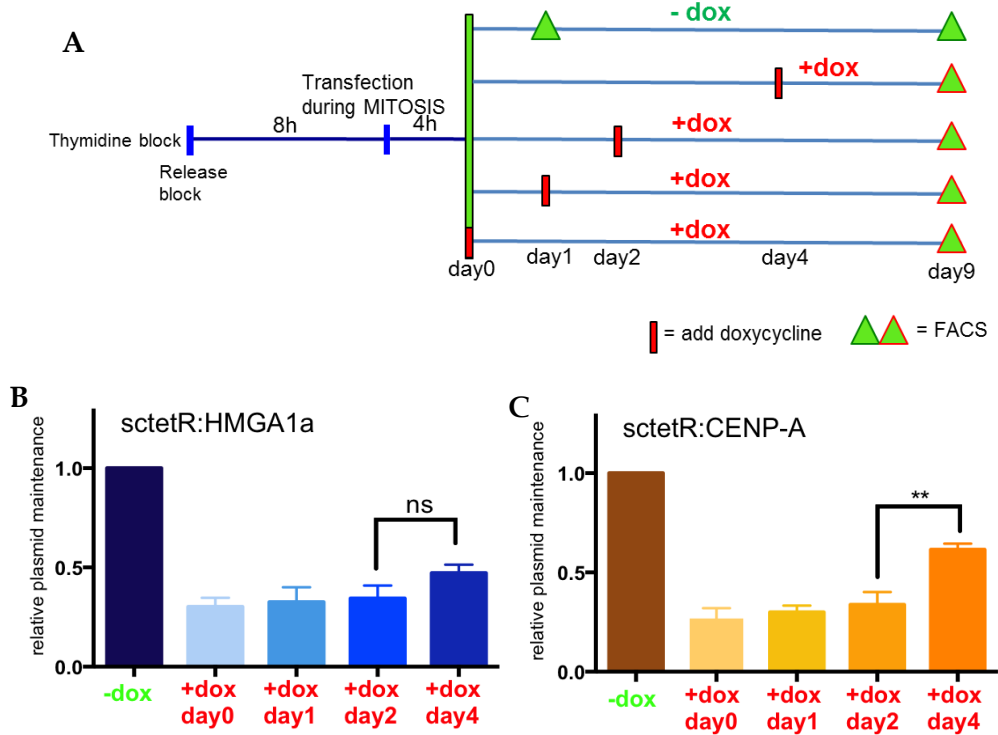
To dissect even shorter centromere maturation time, I transfected synchronized HEK293EBNA1<sup>+</sup> cells with either sctetR:HMGA1a tetO-DS or pCON<sup>CENP-A</sup> plasmids during mitosis and added doxycycline at different time points to repress the binding of tet transactivator to the operator elements (Figure 37 A). Since endogenous CENP-A incorporation is known to take place from late mitosis to G1 phase (McKinley and Cheeseman, 2016), I transfected the cells during mitosis to prevent CENP-A incorporation during the first cell cycle.

HEK cells were released from a thymidine block for 16 hours by washing with PBS and adding cell growth medium. Eight hours after release, when most cells entered mitosis, cells were transfected with the tetO-DS reporter plasmids expressing either sctetR:HMGA1a or sctetR:CENP-A. Doxycycline was added to one plate immediately after transfection (day0) and to other plates one, two or four days later as shown in Figure 37 A. At day1 after transfection the initial amount of transfected plasmids was quantified by FACS measurement of gfp positive cells. On day9 after transfection the plasmid maintenance under different conditions was determined (Figure 37 A). Transfected cells without doxycycline treatment were also analyzed and set as 100 % reference.

The relative plasmid maintenance for the sctetR:HMGA1a plasmids analyzed on day9 after transfection was not significantly different if doxycycline was added on day2 or day4. The slight differences in plasmid maintenance, if day0 and day4 were compared, resulted from the time of doxycycline present in the medium. If doxycycline was added on day0 and analyzed on day9, the plasmids were not established in the cells. If the treatment started at day4, plasmids were initially established and then lost within five days. Contrary to the sctetR:HMGA1a plasmids, the pCON<sup>CENP-A</sup> plasmids were stably maintained in the HEK293EBNA1<sup>+</sup> cells if the sctetR:CENP-A targeting was inhibited at day4. The artificial centromere was established within four days, which led to maintenance of these plasmids in presence

## 4 RESULTS

of doxycycline. Relative plasmid maintenance was significantly reduced if doxycycline was added already two days after transfection or earlier.



**Figure 37: Two days of establishment are not enough**

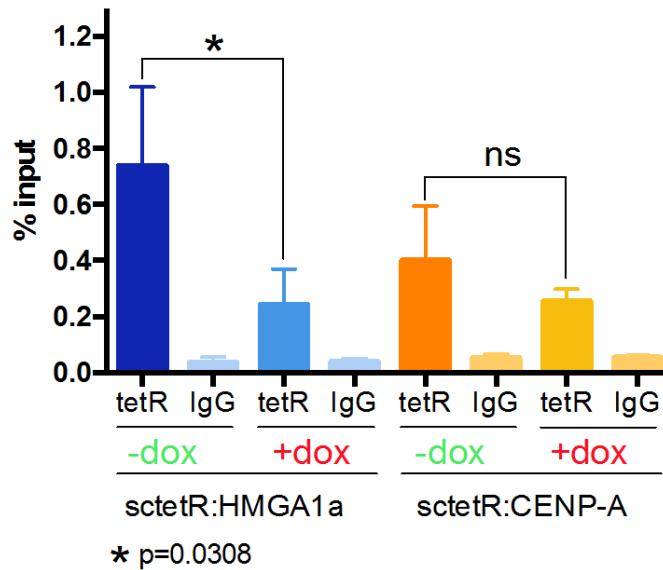
A) Timeline of experimental setup. Cells are transfected during mitosis and then spitted into conditions with and without doxycycline on day0, 1, 2 and 4.

B) tetO-DS sctetR:HMGA1a reporter plasmid loss is not significantly different when doxycycline is added.

C) tetO-DS sctetR:CENP-A is established within four days. If doxycycline is added after two days, plasmids are lost significantly more. (n=3, Mean+SD, \*\*p=0.0074, p-value: unpaired t test with Welch's correction, two-tailed)

These results indicate, that four days are sufficient for the establishment of centromere activity on pCON, whereas two days are not. It is likely that recruitment of CENP-A is independent on CENP-A targeting because of a self-propagation mechanism for CENP-A incorporation. To further verify that, a ChIP experiment in presence and absence of doxycycline was performed. There is no difference in binding of sctetR:CENP-A to tetO sites in both conditions expected. In contrast, the sctetR:HMGA1a binding to tet operators in presence of doxycycline is anticipated to be reduced.

For this tetR ChIP, I transfected HEK293EBNA1<sup>+</sup> cells with either sctetR:HMGA1a tetO-DS or pCON<sup>CENP-A</sup> reporter plasmids. Four days after transfection, I added 2 µg/ml doxycycline for 24 h. After cross-linking in presence of doxycycline a tetR immuno precipitation was performed. Quantitative PCR revealed a 4-fold reduced binding of sctetR:HMGA1a to tetO in presence of doxycycline. In contrast, binding of sctetR:CENP-A to tetO was not influenced by inhibiting the binding capacity of sctetR to tetO with doxycycline (Figure 38).



**Figure 38: ChIP in presence and absence of doxycycline**

Binding of sctetR:HMGA1a to tetO is significantly reduced in presence of doxycycline. Binding of sctetR:CENP-A to tetO is not significantly affected by doxycycline. (n=4, Mean+SD, p-value: unpaired t test with Welch's correction, two-tailed)

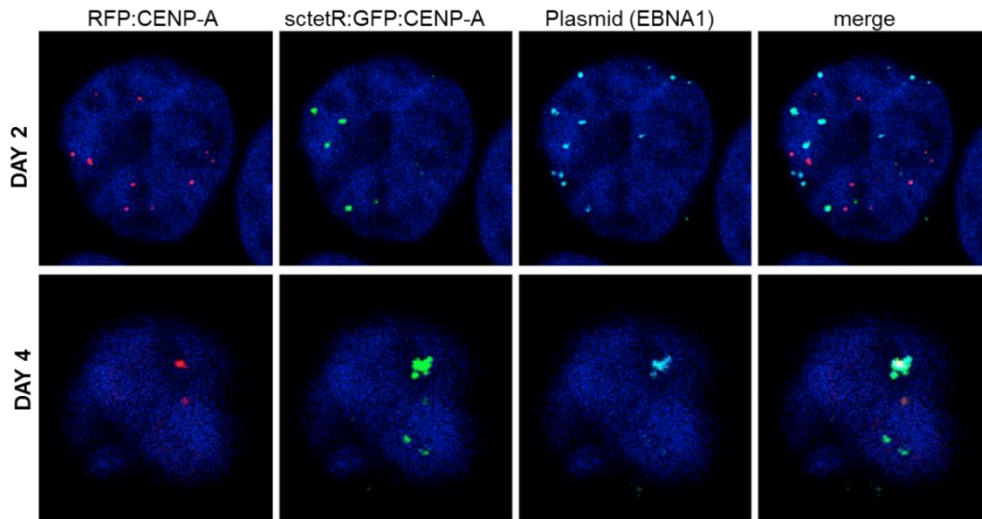
It is confirmed that within four days the centromere inheritance is established on the pCON<sup>CENP-A</sup> plasmids (Figure 37), whereas two days are not sufficient for centromere establishment. Maintenance of neo-centromeres is independent of CENP-A targeting, demonstrated by ChIP of tetR fusion proteins. In addition, the sctetR:CENP-A fusion proteins are incorporated into centromeric nucleosomes (Figure 34). The hypothesis resulting from these observations is that endogenous CENP-A is also recruited to the

plasmids and this leads to the targeting-independent and self-propagating centromere inheritance on pCON<sup>CENP-A</sup>.

#### 4.7 Recruitment of CENP-A

To test the hypothesis, if endogenous CENP-A is recruited to pCON<sup>CENP-A</sup> centromeres after establishment, I performed immune fluorescence microscopy. The endogenous and targeting-independent CENP-A protein was represented by a RFP:CENP-A fusion protein.

I transfected HEK293EBNA1<sup>+</sup> cells with the *scetR*:GFP:CENP-A expressing pCON<sup>CENP-A</sup> reporter plasmid and analyzed the recruitment of a RFP:CENP-A protein two and four days after transfection by co-staining of plasmid signals and RFP:CENP-A. According to the result depicted in Figure 37 a recruitment of endogenous CENP-A, represented by RFP:CENP-A, was expected on day4, but not on day2.



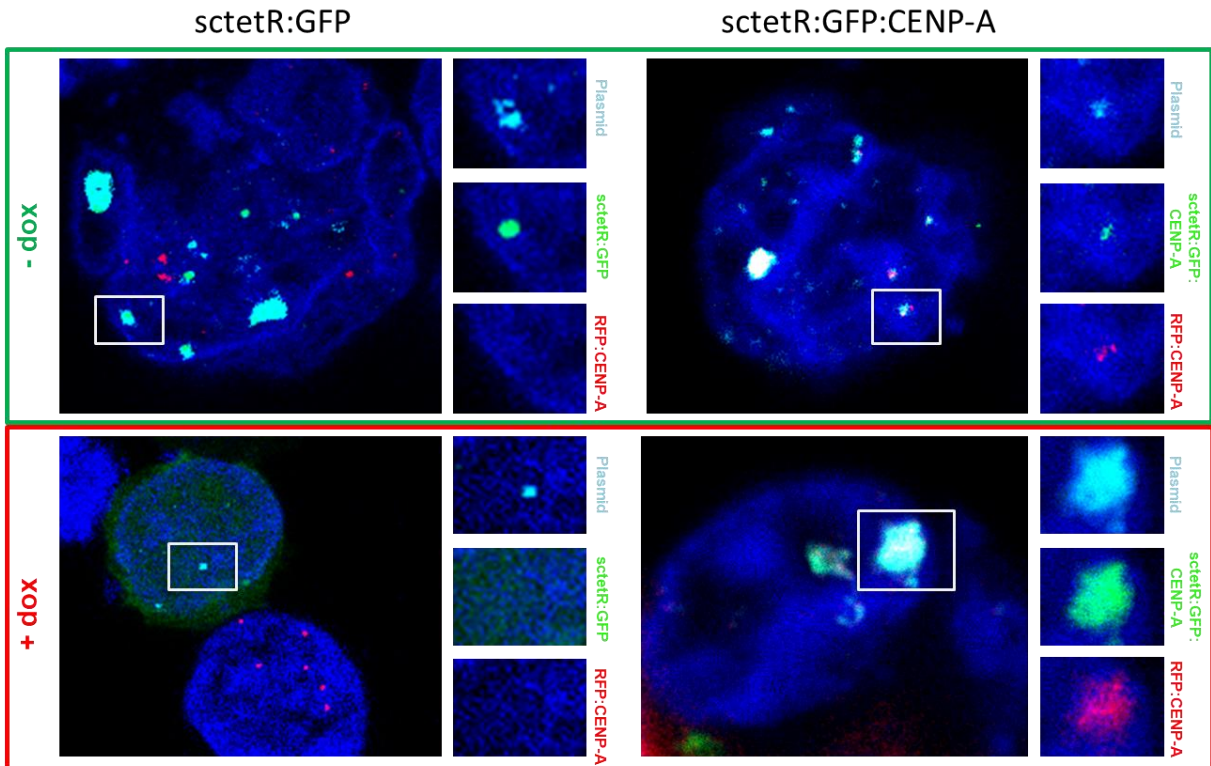
**Figure 39: Recruitment of RFP:CENP-A to plasmids**

RFP:CENP-A co-localizes with *scetR*:GFP:CENP-A and EBNA1 on day 4 after transfection. 2 days after transfection RFP:CENP-A is not recruited to plasmids.



Indeed, what was observed in this immune fluorescence analysis was, that RFP:CENP-A was recruited to the plasmids four days, but not two days after transfection. Two days of centromere maturation were not sufficient to incorporate endogenous CENP-A (represented by RFP:CENP-A) into plasmid chromatin, whereas after four days RFP:CENP-A was loaded into the artificial centromeres on the plasmids (Figure 39).

In order to investigate if recruitment of RFP:CENP-A to the plasmids is independent on the targeting of *sctetR*:GFP:CENP-A, doxycycline was added to the cells. HEK293EBNA1<sup>+</sup> cells were transfected with either *sctetR*:GFP:CENP-A expressing pCON<sup>CENP-A</sup> plasmids or, as reference, with *sctetR*:GFP expressing tetO-DS reporter plasmids and incubated for four days. Then, the RFP:CENP-A expression plasmid was transfected and cells were split into two conditions: presence and absence of 2 µg/ml doxycycline.



**Figure 40: Recruitment of RFP:CENP-A to plasmids in presence of doxycycline**

RFP:CENP-A co-localizes with tetO-DS plasmids in presence and absence of doxycycline when *sctetR*:GFP:CENP-A is targeted. RFP:CENP-A is not recruited when *sctetR*:GFP is targeted. In presence of doxycycline *sctetR*:GFP is not co-localized to plasmids, whereas *sctetR*:GFP is co-localized with plasmids when no doxycycline is present.

As already demonstrated in Figure 39, RFP:CENP-A was recruited to plasmids in absence of doxycycline. In addition a co-staining of RFP:CENP-A, sctetR:GFP:CENP-A and EBNA1 in presence of doxycycline was detected. By highlighting this observation, it indicates that not only endogenous or RFP:CENP-A but also sctetR:GFP:CENP-A is incorporated into pCON<sup>CENP-A</sup> by endogenous mechanisms.

In contrast, initial targeting of sctetR:GFP did not recruit RFP:CENP-A to tetO-DS reporter plasmids. The control experiment also illustrated that applying doxycycline led to decreased tetR-tetO interaction, because the sctetR:GFP signal was distributed in the whole nucleus after adding doxycycline, whereas it co-localized with EBNA1, representing plasmids, in absence of doxycycline (Figure 40).

In summary, we found out that after establishment of centromere identity on plasmids, these were maintained stable and independent on CENP-A targeting. Recruitment of kinetochore proteins led to the active segregation of the plasmids during mitosis. In addition my results showed that it takes only four days to establish an inheritable centromere on plasmids in human cells. All these findings indicate that the inheritable centromere establishes epigenetic modifications to self-propagate and maintain its identity.

Therefore, the next aim was to reveal the epigenetic characteristics of the plasmid centromeric chromatin that lead to its inheritance. For that purpose, I utilized the CRISPR/Cas9 targeting system, introduced in chapter 4.3.1, to purify plasmids with matured centromeres out of the human cells. After purification of the plasmids, several histone modifications were analyzed, since these are potential candidates for epigenetic inheritance of centromere identity.

### 4.8 CRISPR/Cas9 targeting system to analyze histone modifications on the artificial centromere

To purify plasmid centromeres, established after CENP-A targeting, and analyze their epigenetic histone modification pattern, I needed a system to isolate the plasmid DNA from human cells. The master student Alejandro Freyermuth compared a LexA dependent targeting system, where a LexA:TAP fusion protein was targeted to LexA sites on plasmids with an innovative CRISPR/Cas9-dependent purification method during his master thesis project.

CRISPR/Cas9 is a new tool to target and modify DNA systematically and easily. It is a bacterial system altered and engineered to use also in mammalian cells. Cas9 is an endonuclease that cuts double stranded DNA after specific targeting by a guide RNA. This guide RNA is complementary to the DNA sequence where Cas9 should cut the DNA. In the engineered system the nuclease activity leads to homologous recombination of the DNA strands and results in either deletion of a DNA fragment or in the insertion or replacement by co-transfecting a target DNA-sequence (Sander and Joung, 2014).

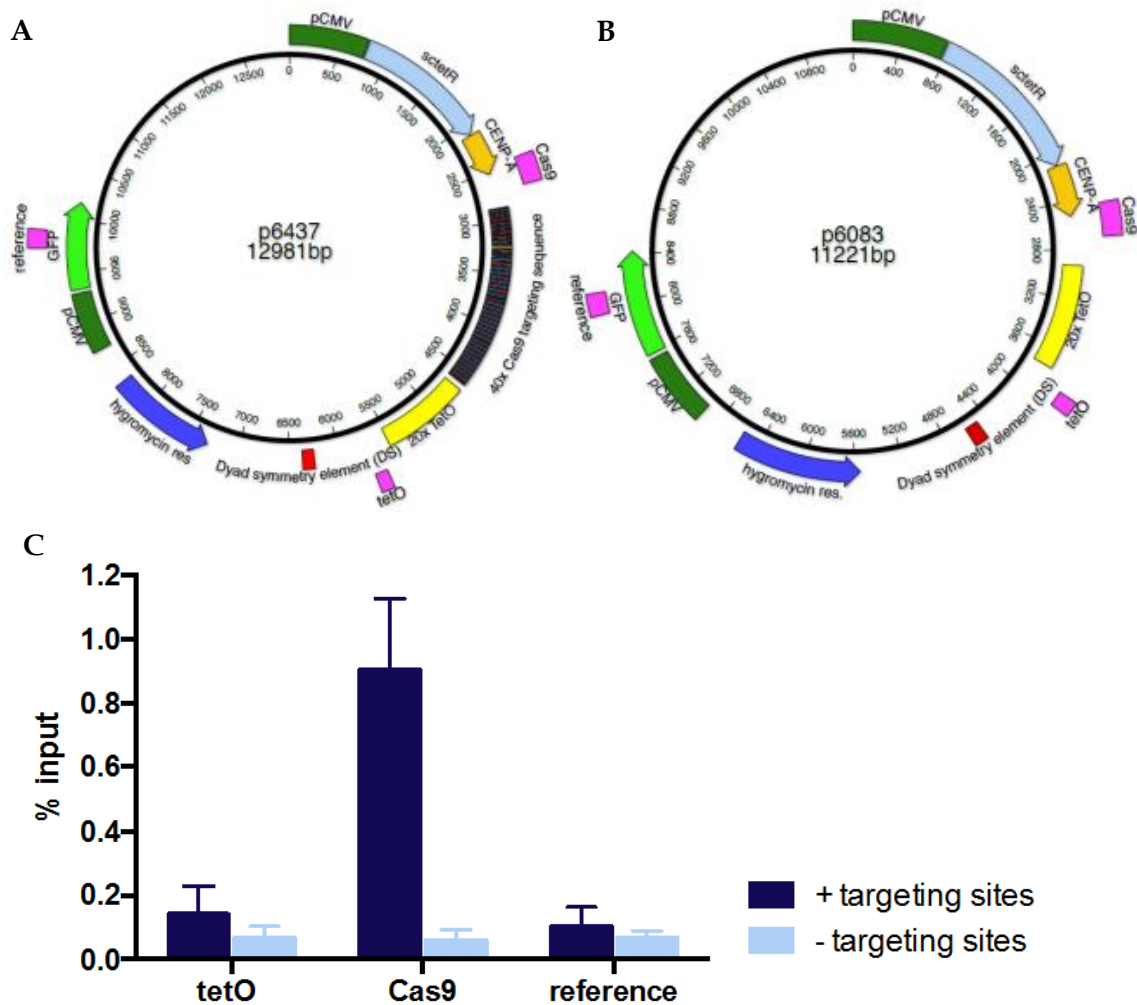
In our system it turned out that the CRISPR/Cas9 targeting method was more specific than the LexA system. We used the nuclease deficient Cas9 protein that was tagged with a tandem affinity purification (TAP) tag, consisting of a twin-STREP tag and a Flag tag (see Figure 17, chapter 3.2.1.1).

Since this Cas9 dependent plasmid purification system was not used before, we characterized it in its targeting specificity to plasmids and optimized a purification protocol to analyze epigenetic histone modifications established on the plasmid by targeting CENP-A.

#### 4.8.1 Specific binding of Cas9 to the targeting sites in the plasmid system

In order to investigate binding efficiency of Cas9:mCherry:TAP to the targeting sites on the reporter plasmids, I performed a STREP pull down according to the ChIP protocol (see chapter 3.2.7).

HEK293EBNA1<sup>+</sup> cells, stable expressing Cas9:mCherry:TAP, were transfected with the pCON<sup>CENP-A</sup> plasmid with and without a repeat of 40 Cas9 targeting sites. Readout of this pull down was a qPCR with primer locations several (kilo) bases away from the targeting sites (Figure 41 C tetO; reference) and in close proximity to the targeting sites (Figure 41 C Cas9).



**Figure 41: ChIP of plasmid with and without targeting sites**

A) Representation of tetO-DS scd4R:CENP-A reporter plasmid with Cas9 targeting sites.

B) Representaion of tetO-DS scd4R:CENP-A reporter plasmid

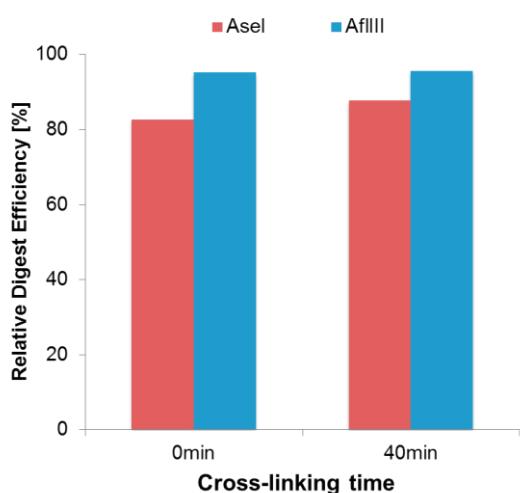
C) qPCR of plasmid sites located at tetO and Cas9 sites. Targeting of Cas9:TAP is highly specific to Cas9 targeting sites, if those are present on reporter plasmid. (n=2, Mean+SD)

Specific binding of the Cas9:mCherry:TAP fusion protein was observed after STREP ChIP. Significant enrichment of Cas9 was found at the Cas9 targeting sites on the plasmid that contains these sites, whereas no Cas9 is bound to reference regions or to the plasmid without any specific targeting sites (Figure 41 C).

Hence, this system is suitable for the specific purification of pCON<sup>CENP-A</sup> plasmids by a STREP pull down and to further characterize the epigenetic information established at the artificial centromere on the plasmids.

### 4.8.2 High digest efficiency of plasmid DNA

For the analysis of the matured centromeres only the region within and around the tetO array is of interest. However, the pCON<sup>CENP-A</sup> including the Cas9 targeting sites has a size of around 13 kbp. The fragment of interest, with the matured centromere, is around 3 kbp in size.



**Figure 42: High digest efficiency in un-cross-linked and heavily cross-linked sample**

*Asel* digest (red bars) is a little less efficient than *AflIII* digest (blue bars).

Overall digest efficiency of plasmid DNA is very high with more than 80 %

Data: Alejandro Freyermuth (n=1)

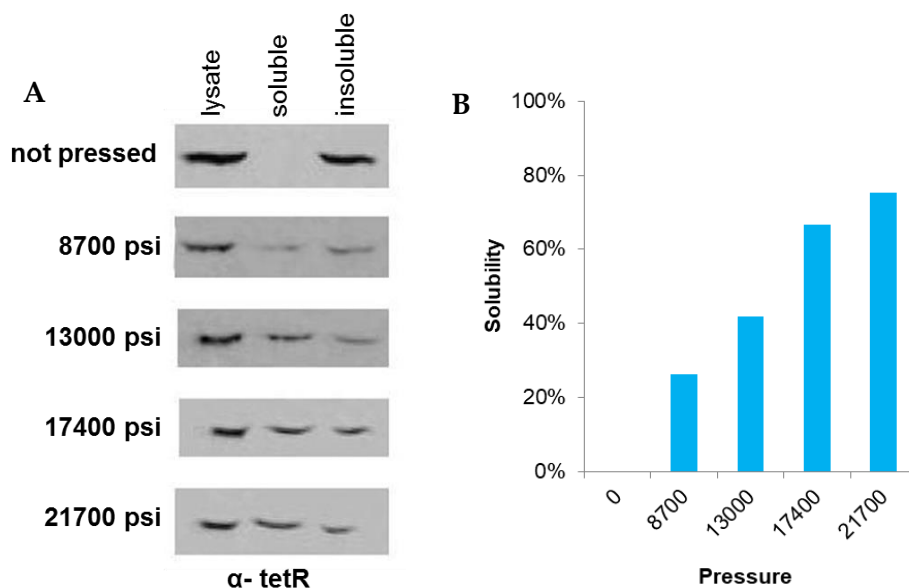
In order to reduce background signals, induced by the remaining plasmid, the plasmid DNA was digested by specific endonucleases. Combining an *Asel* and *AflIII* restriction enzyme digest of whole cell extract transfected with reporter plasmids led to fragmentation of genomic DNA and to site specific digest of plasmid DNA. The resulting fragment included centromeric nucleosomes at the tetO sites and the Cas9 targeting sites for purification (Figure 18, chapter 3.2.9.3).

To verify the digest efficiency of *AseI* and *AflIII* on the plasmids a quantification by qPCR was performed by the master student Alejandro Freyermuth under my supervision. He transfected HEK293EBNBA1<sup>+</sup> cells with pCON<sup>CENP-A</sup> plasmids containing the Cas9 targeting sites. Four to six days later, he prepared cell extracts either not or heavily (40 min) cross-linked. After double-digestion with *AseI* and *AflIII*, he extracted the DNA and quantified digest efficiencies by qPCR according to Hagege et al. (equation see chapter 3.2.8) (Hagege et al., 2007).

The primer pairs covering the *AseI* or *AflIII* digest region included one restriction site of *AseI* or *AflIII* respectively. *AflIII* restriction digest showed a slightly better digest efficiency depicted in Figure 42 (blue bars). However, in general a high digest efficiency of 80 % for *AseI* and 95 % for *AflIII* was observed in un-cross-linked and cross-linked chromatin. These efficiencies were suitable to continue the purification protocol to investigate histone modifications established on the artificial centromere.

### 4.8.3 Re-Solubilization by high pressure in a French pressure cell press

To remove insoluble proteins and protein aggregates, which would interfere with the STREP-beads during purification, Alejandro centrifuged the cell lysate. By doing that, we observed that almost all material was pelleted and no proteins remained soluble in the supernatant after centrifugation (Figure 43, not pressed). Hence, we decided to solubilize the cell lysate by disposing high pressure mediated by a French pressure cell press, like used in the PICH (Proteomics of Isolated Chromatin) protocol on the ribosomal RNA gene promoter developed in J. Déjardins group (Ide and Dejardin, 2015).



**Figure 43: Protein re-solubilization by French Press**

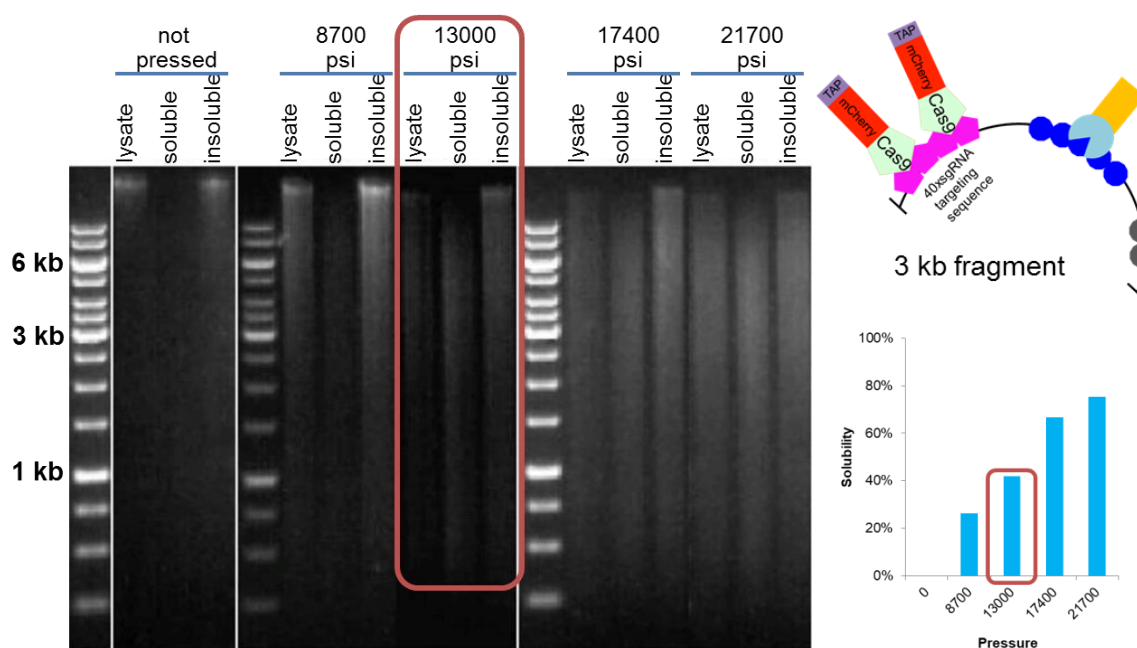
A) Re-solubilization of proteins. Without French Press proteins are not soluble, with increasing pressure the protein solubilization is also rising.

B) Representation of quantification of protein re-solubility.

Data: Alejandro Freyermuth, *Master thesis*

As expected the protein solubility, as well as DNA solubility increased with rising pressure (see Figure 43 A for protein and Figure 44 for DNA). At a pressure of 8700 psi only 20 % of protein was re-solubilized, whereas at a pressure of 21700 psi almost 80 % of protein was soluble (Figure 43 B). To define the optimal pressure for re-solubilization of chromatin, one also has to take care about the DNA re-solubilization and the fragmentation of DNA occurring by shearing in the French pressure cell press.

Without pressing the lysate and loading extracted DNA on an agarose gel, a distinct band above the 10 kb marker was observed (Figure 44; not pressed, lysate). This DNA was insoluble (Figure 44; not pressed, insoluble) and therefore not accessible in our purification approach. After exposing the lysate to high pressure, the DNA was re-solubilized but also fragmented (Figure 44; 13000 psi-21700 psi). To ensure to have enough re-solubilization but not too much shearing to break the 3 kb plasmid fragment we were aiming to analyze, we used a pressure of 13000 psi (Figure 44 marked by red rectangle). In this setting the DNA size of bulk soluble DNA was between 3 kb and 6 kb and the solubility was at 40 %.



**Figure 44: DNA re-solubilization by French Press**

DNA is re-solubilized but also fragmented by French Press. To pull down artificial centromere at tetO sites, fragment size needs to be more than 3 kb. At 13000 psi conditions are suitable for purification of plasmids. Fragment size is more than 3 kb and re-solubilization is around 40 %.

Data: Alejandro Freyermuth, *Master thesis*

The conditions for STREP pull down were established as the following: 5 minutes cross-link with 1 % formaldehyde; 13000 psi French press; 2 h digest with *AseI* and *AflIII*.

By performing a STERP purification of Cas9:mCherry:TAP under these conditions, I investigated the histone modifications that establish on neo-centromeres after CENP-A targeting.

#### 4.8.4 Efficient STREP pull-down of Cas9:mCherry:TAP fusion protein

With the Cas9 dependent STERP purification I analyzed the composition of a matured artificial centromere on the pCON<sup>CENP-A</sup> plasmid. We knew from previous experiments, described in chapter 4.6.3, that the centromere identity is established



already after four days. However, for technical reasons I analyzed modifications, possibly dictating the centromere inheritance, after a longer establishment time.

Thus, I transfected HEK293EBNA1<sup>+</sup> cells with the pCON<sup>CENP-A</sup> reporter plasmid either including Cas9 targeting sites or without Cas9 targeting sites. Both plasmids were stable established in the cells for three weeks under selection pressure. After re-transfecting these cells with the Cas9:mCherry:TAP expression plasmid together with the sgRNA encoding plasmid and incubating them for additional five days, they were harvested and cross-linked. The readout of the STREP pull down after French Press and restriction digest was performed by Western Blot.



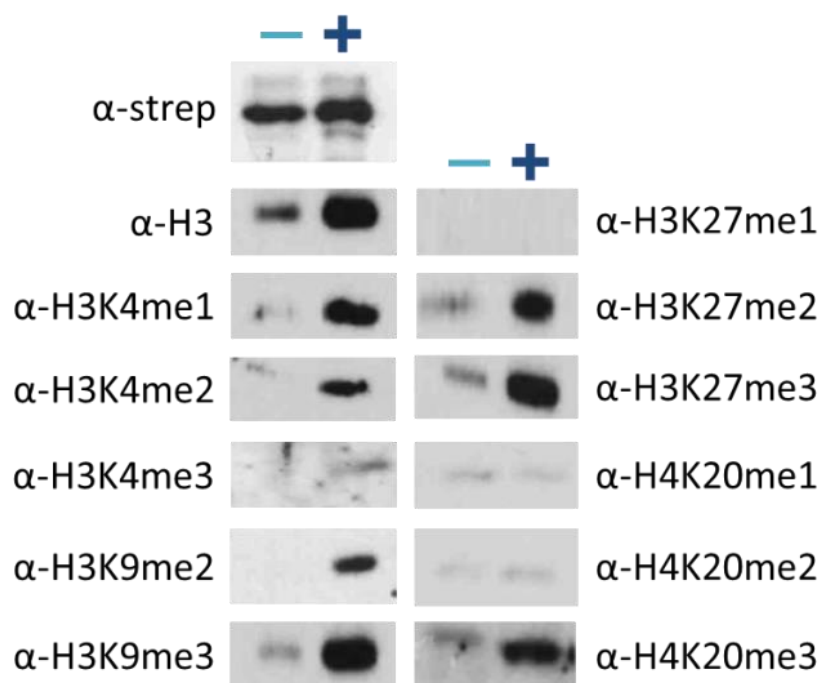
**Figure 45: STREP pull down after French press and digest**

Cas9:mCherry:TAP Western Blot for analysis of solubility and pull down efficiency with STREP beads. Before French Press almost no protein is soluble. After French Press and digest around 50 % proteins are soluble. This is the case for cells transfected with plasmids with and without targeting sites.

As verified in the Western Blot (Figure 45), the Cas9:mCherry:TAP protein became soluble after exerting a pressure of 13000 psi by French press. The same amount of protein stayed soluble after restriction enzyme digest. The soluble fraction was then used for STREP pull down. Besides there was a lot unbound Cas9:mCherry:TAP remaining on the beads, the elution of the protein was sufficient to investigate co-precipitation of histones and modified histones established on plasmid chromatin.

#### 4.8.5 Co-purification of modified histones

In order to determine which histones and possible histone modifications are present around tet operator sites after the establishment of an artificial centromere, different Western Blot analyses were performed after the STREP pull down depicted in Figure 45. Equal amounts of eluted proteins of the STREP pull down of cell extracts containing the pCON<sup>CENP-A</sup> reporter plasmids with targeting sites were compared with extracts containing the reporter plasmids without the Cas9 targeting sites.



**Figure 46: Co-purification of histones and specific modifications**

Elutions of STREP pull down analyzed in co-precipitation of histone H3 and histone H4 modifications. Analyzed were plasmids without (-) and with (+) Cas9 targeting sites. (n=1)

The pull down of Cas9:mCherry:TAP proteins in cells containing the pCON<sup>CENP-A</sup> with targeting sites (+) and without targeting sites (-) was comparable (Figure 46, α-strep). However, the co-precipitation of histones and histones carrying specific modifications was higher when Cas9 targeting sites were present on the reporter plasmid (compare Figure 46 α-H3). Most prominent modifications were H3K9me3, H3K4me1 and H3K4me2. In addition H3K9me2, H3K27me2 and H3K27me3 were

enriched on plasmid chromatin after CENP-A targeting. Histone H4 showed a trimethylation state on lysine 20. No significant signal compared to pCON<sup>CENP-A</sup> without Cas9 targeting sites was observed for H4K20me1, H4H20me2 and H3K4me3. For H3K27me1, there was no signal obtained at all (Figure 46,  $\alpha$ -H3K27me1; input analysis not shown).

Most of these histone modifications, especially H3K4me2 and H3K9me3, are known to be present on human centromeric and pericentromeric chromatin (McKinley and Cheeseman, 2016). Finding these also in the plasmid system gives a hint, that initial targeting of CENP-A to a foreign DNA locus leads to the maturation of an inheritable centromere identity that is comparable to endogenous centromeres.

## 5 DISCUSSION

I used the pCON<sup>CENP-A</sup> plasmid system as a suitable tool to analyze *de novo* centromere establishment and epigenetic maturation of artificial centromeres. Site-specific targeting of CENP-A leads to the establishment of an inheritable centromere identity. With the pCON<sup>CENP-A</sup> plasmid system, I addressed three distinct aspects:

I) The CENP-A dependent plasmid segregation mechanism by live cell imaging of cells containing pCON<sup>CENP-A</sup>. Furthermore, I analyzed their capacity to recruit kinetochore components by immune fluorescence.

II) The minimal plasmid establishment time required for long-term maintenance and CENP-A self-propagation. Plasmid maintenance became independent of CENP-A targeting and I examined the minimal timeframe of centromere maturation in which centromeres begin to self-propagate.

III) The histone modifications present on matured plasmid centromeres. Therefore, I developed a Cas9-dependent targeting system and purified plasmids with matured centromeres from human cells.

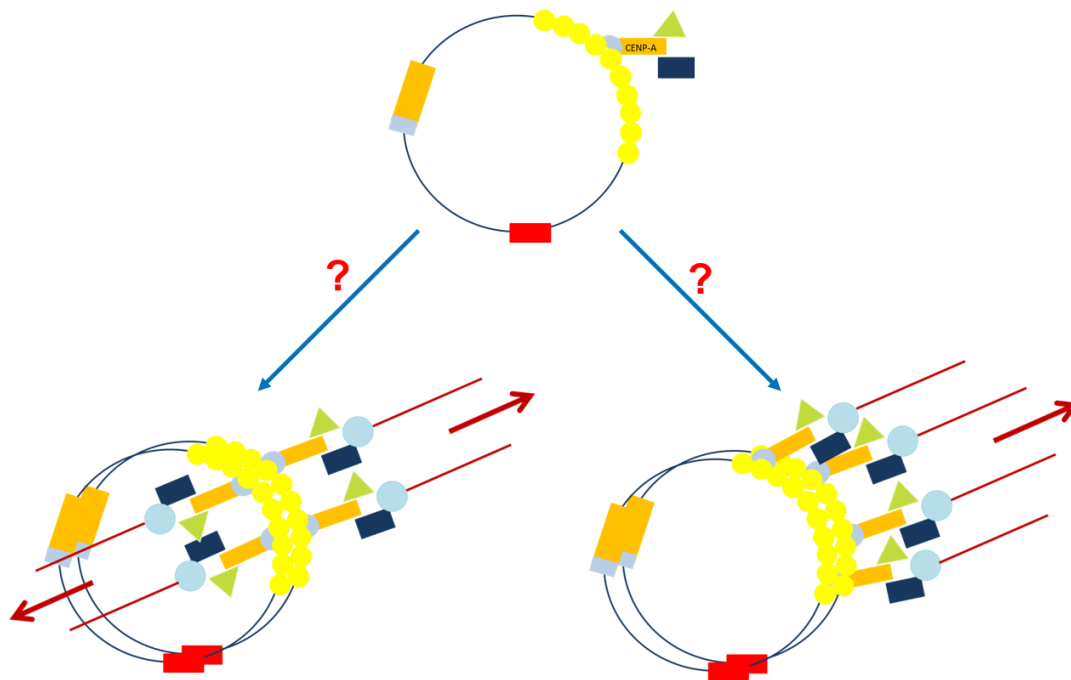
### 5.1 Plasmid segregation mechanism after CENP-A targeting

Initial targeting of sctetR:CENP-A to a tetO array of pCON<sup>CENP-A</sup> leads stable maintenance of pCON<sup>CENP-A</sup> plasmids within human cells (Figure 14), because of the formation of a functional neo-centromere on these plasmids. At endogenous chromosomes centromeres are the platform for kinetochore assembly and they mediate interaction of microtubules. Thus they regulate active segregation during mitosis. Since centromeres are formed on the plasmids as well, we hypothesize an active segregation mechanism for pCON<sup>CENP-A</sup> plasmids, like it is observed for endogenous chromosomes.

To address that question, we investigated the capacity of pCON<sup>CENP-A</sup> to recruit kinetochore proteins. I chose the inner kinetochore protein CENP-C and the outer kinetochore protein Ndc80 as representative components. CENP-C is one of the most important inner kinetochore proteins, because it directly interacts with the CENP-A nucleosome and stabilizes it (Falk et al., 2015). In addition CENP-C plays a role in self-propagation of centromere inheritance (McKinley and Cheeseman, 2016). Presence of CENP-C at artificial centromeres on plasmids already indicates that these are able to self-propagate and recruit CENP-A in a cell cycle dependent manner. On this account, I examined the recruitment of CENP-C to artificial neo-centromeres on the plasmids by immune fluorescence. The targeting of scetR:GFP:CENP-A results in highly enriched localization of CENP-C signals at the plasmids compared to targeting of scetR:GFP (Figure 31). 55 % of pCON<sup>CENP-A</sup> plasmid signals, represented by EBNA1 staining, overlap with CENP-C spots in a scetR:GFP:CENP-A targeting dependent manner. If scetR:GFP only is targeted, no overlap of CENP-C with EBNA1 is observed. This indicates that CENP-C recruitment is dependent on CENP-A targeting and that the neo-centromeres are able to attract CCAN components. CCAN is constitutively bound to centromeres and provides a binding site for outer kinetochore proteins during mitosis. Ndc80 is the major sub-complex for active segregation during mitosis, because its HEC1 subunit directly interacts with spindle microtubules (Wei et al., 2007). Only if Ndc80 is recruited to plasmids, these will be able to actively segregate by microtubule interaction. I analyzed presence of HEC1 at neo-centromeres, also by immune fluorescence, in mitotic cells. Like for the CENP-C protein, recruitment of HEC1 is dependent on initial CENP-A targeting. Overlap between plasmid signals and HEC1 signals is with 60 % significantly higher if scetR:GFP:CENP-A was targeted compared to scetR:GFP targeting (Figure 31). These data suggest that neo-centromeres on pCON<sup>CENP-A</sup> are capable to recruit kinetochore components. However, the segregation mechanism is not resolved with these approaches.

Immune fluorescence experiments in the *Drosophila melanogaster* system demonstrated the association of plasmid neo-centromeres to spindle microtubules

(Figure 11, in collaboration with P. Heun). We also observed that segregation of these plasmids is not timely regulated as segregation of chromosomes (Mendiburo et al., 2011). Two models of plasmid segregation are conceivable, as represented in Figure 47.



**Figure 47: Segregation mechanism of pCON<sup>CENP-A</sup> plasmids**

After replication plasmids are segregated by recruitment of spindle microtubules during mitosis. Plasmids can be attached at both sides and segregated equally to daughter cells like endogenous chromosomes (left). Or plasmids are attached by microtubules from just one direction and pulled towards one spindle pole asymmetrically (right).

First, plasmids are either bound by microtubules at two sites of centromere region, like in endogenous chromosomes (Figure 47, left). This leads to segregation of replicated plasmids and equal distribution to daughter cells. Alternatively, replicated plasmids are attached to microtubules from one direction and are segregated asymmetrically to the spindle poles (Figure 47, right).

To clarify the segregation mechanism of plasmids that contain an artificial centromere, I performed live cell imaging experiments. I generated a CRISPR/Cas9 dependent targeting system to specifically visualize pCON<sup>CENP-A</sup> plasmids in living cells. The Cas9:3xmCherry protein approach was already established in multicolor imaging of chromosomal loci (Ma et al., 2015). By counterstaining cells with a live cell

tubulin dye, I aimed to detect microtubule-plasmid interaction during mitosis in living cells.

Because of technical problems, I have not been able to finalize this project. In initial experiments, I followed one cell over the cell cycle and detected the Cas9:3xmCherry signal during cell division. In this experiment one plasmid spot localized towards just one spindle pole (Figure 28), indicating an asymmetric plasmid segregation for this cell.

In conclusion, the molecular segregation mechanism of pCON<sup>CENP-A</sup> with neo-centromeres is still unclear. Fly experiments indicate that segregation of plasmids might be earlier than segregation of endogenous chromosomes because plasmids are detected at spindle poles already at metaphase (Mendiburo et al., 2011). It is likely that plasmid segregation is not part of the spindle assembly checkpoint (SAC). Cells detect correct attachment of microtubules to both sites of chromosomes by generation of tension and tension signaling by kinetochores at the SAC. Only if kinetochores of all chromosomes detect tension, generated by cohesin proteins, mitotic progression takes place (Kim and Yu, 2015). In replicated plasmids, the centromeres are most likely not arranged as cohesin attached pairs, a prerequisite for the spindle checkpoint. It is well possible that plasmids are distributed asymmetrically in human cells (Figure 28).

The composition of the minimal kinetochore complex and proteins recruited to plasmids once centromere is established, like cohesins, could be determined by a tandem affinity purification and subsequent mass spectrometric analysis. In contrast to *in vitro* assembly of purified proteins, like in the study of Guse et al. (Guse et al., 2011), the pCON<sup>CENP-A</sup> system and its analysis after purification would reveal an *in vivo* assembled minimal kinetochore and might allow structural characterization of this huge protein complex by AFM (atomic force microscopy) studies.

## 5.2 Establishment of self-propagating centromeres on plasmids

In the pCON<sup>CENP-A</sup> reporter plasmid, the sc tetR:CENP-A fusion protein is targeted to plasmids by tetR-tetO interaction. The transactivator tetR is bound to the specific tetO sequence by dimerization of two tetR proteins. In the sc tetR fusion protein both dimers are connected by a linker. Binding of tetR to tetO can be inhibited with doxycycline by changing tetR dimer structure that the interaction to tetO is interrupted.

Site-specific targeting of sc tetR:CENP-A was investigated by ChIP and immune fluorescence. Both techniques verified the site specific targeting of sc tetR:CENP-A to plasmids (Figure 20 and Figure 22). Another important finding of the pCON<sup>CENP-A</sup> system was its stability in human cells. These plasmids showed long-term maintenance without selective pressure for five months (Figure 19). In collaboration we already reported, that targeting of CID, *Drosophila* CENP-A, is sufficient for establishment of an inheritable centromere identity in fly cells (Mendiburo et al., 2011). With pCON<sup>CENP-A</sup> we observed that also this plasmid maintenance becomes independent of CENP-A targeting after a centromere has established on the plasmids (Figure 36). Hence, we presume an artificial neo-centromere established on the plasmids, that is self-propagating and regulates its CENP-A incorporation.

In order to test this hypothesis, different experimental approaches were used. First hints about CENP-A targeting independent centromere inheritance were given by ChIP of sc tetR:CENP-A at two different time points. Precipitation of sc tetR:CENP-A after four days establishment revealed localization of the fusion protein at tet operator sites. In contrast, at the matured centromere after three weeks establishment time, no sc tetR:CENP-A was detectable at the plasmid tetO array. To further validate targeting-dependency of centromere inheritance we used plasmid maintenance as readout. After an establishment time of seven days, we added doxycycline into the system and investigated pCON<sup>CENP-A</sup> maintenance when tetR-tetO interaction is inhibited (Figure 36). Even when the targeting of sc tetR:CENP-A to the plasmids is interrupted, these were maintained several weeks. This implied a centromere



inheritance mechanism that is independent of CENP-A targeting. We already knew from *Drosophila melanogaster* experiments, that an epigenetic mark is established at artificial centromeres, which leads to its self-propagation and incorporation of additional CENP-A (Mendiburo et al., 2011). For that reason we analyzed recruitment of CENP-A, that was not mediated by the tetR transactivator. RFP:CENP-A was a representative for endogenous CENP-A, potentially recruited to the neo-centromeres. Two days after introducing pCON<sup>CENP-A</sup> into human cells, no RFP:CENP-A is incorporated into plasmid centromeres. In contrast, two days later, four days after transfection, RFP:CENP-A was present at pCON<sup>CENP-A</sup>, together with sctetR:GFP:CENP-A (Figure 39). In an additional experiment we revealed that not only RFP:CENP-A but also sctetR:GFP:CENP-A is recruited to plasmid by a targeting independent mechanism (Figure 40).

These results demonstrate that CENP-A targeting is sufficient to establish an inheritable centromere identity on plasmids in the human system. These centromeres are self-propagating and regulate CENP-A incorporation.

However, how the of CENP-A containing neo-centromere is restricted and how many CENP-A nucleosomes are needed for the establishment of an artificial centromere was not completely solved with our approaches. For endogenous centromeres it is known that expression of exogenous CENP-A leads to a downregulation of endogenous CENP-A in human cells (Jansen et al., 2007) and it was observed that CENP-A overexpression leads to spreading of CENP-A to chromosome arms (Gascoigne et al., 2011). This indicates that CENP-A restriction to centromere core is partly determined by CENP-A levels in the cells and the amount of CENP-A nucleosomes incorporated into chromatin. If threshold level of CENP-A nucleosomes is too low, for example at mis-incorporation in chromosome arms, CENP-A is evicted again. The minimum functional core for epigenetic inheritance of eukaryotic centromeres is 30 to 70 kb of alphoid DNA arrays (Okamoto et al., 2007). For the pCON<sup>CENP-A</sup> plasmid system only 6 CENP-A nucleosomes are sufficient for establishment of centromere inheritance, whereas 6 CENP-A nucleosomes within endogenous chromosomes are not sufficient. We conclude for pCON<sup>CENP-A</sup>, that only

a small number of CENP-A nucleosomes leads to the establishment of a neo-centromere and that each CENP-A triggers new CENP-A incorporation only once, like in endogenous centromeres (Ross et al., 2016).

### 5.3 Four days are sufficient for *de novo* Centromere inheritance

The pCON<sup>CENP-A</sup> establish a stable maintenance in human cells without selective pressure (Figure 36) and targeting of CENP-A is sufficient to create an inheritable centromere identity on plasmids. In addition, already four days after transfection, RFP:CENP-A, is recruited to neo-centromeres by a targeting independent mechanism (Figure 39). The question that arises from these observations is:

How fast is the inheritable centromere identity established on the plasmids?

We knew from previous experiments that seven days of establishment lead to stable plasmid maintenance. Since already after four days RFP:CENP-A is present at artificial centromeres, we investigated plasmid maintenance after an establishment time of four days. We observed that the pCON<sup>CENP-A</sup> plasmids were as stable as after seven days when doxycycline inhibited tetR targeting four days after transfection (Appendix Figure 5). In conclusion this meant that only three to four cell cycles were sufficient for *de novo* formation of neo-centromeres. To ascertain if centromere formation on plasmids is also possible in less than four days, I further shortened the establishment time. CENP-A targeting was inhibited immediately (day0), one day, two days and four days after transfection (Figure 37 A) and plasmid maintenance after nine days was used as readout for centromere formation. Two days of establishment are not sufficient for *de novo* centromere maturation and plasmid maintenance (Figure 37 C), as indicated as well in the RFP:CENP-A immune fluorescence experiment (Figure 39).

For maturation of centromeres, plasmids have to be chromatinized first. The process of chromatinization of plasmids was shown already *in vivo* (Riu et al., 2007). Since the

pCON<sup>CENP-A</sup> encodes the sctetR:CENP-A fusion protein, it also takes some time to express this fusion protein from plasmids, successfully target it and incorporate it into plasmid nucleosomes. After chromatinization, successful targeting of sctetR:CENP-A and its incorporation the centromeres are able to self-propagate its own CENP-A incorporation. The timeframe of these events taking place, which we identified in our experiments, is four days. Four days correspond to three to four cell cycles. We also demonstrated in an immune fluorescence approach that RFP:CENP-A and sctetR:GFP:CENP-A are recruited to plasmid centromeres in presence of doxycycline (Figure 40). This verifies that centromere inheritance and incorporation of CENP-A in plasmids is independent of targeting. In conclusion, we revealed that a short pulse of CENP-A targeting of only three to four cell cycles is sufficient to establish a self-propagating and inheritable centromere on pCON<sup>CENP-A</sup>.

However, our approach did not uncover the mechanism behind CENP-A incorporation and how centromere identity is determined on the plasmids. Since RFP:CENP-A or sctetR:CENP-A in presence of doxycycline are independent of the tetR targeting mechanism, it is likely that CENP-A incorporation into plasmids is mediated by HJURP. HJURP recognizes CENP-A present at artificial centromeres on plasmids and targets and incorporates new CENP-A via the endogenous CENP-A incorporation cycle. HJURP dependent CENP-A incorporation is dependent on the CATD domain of CENP-A (Bassett et al., 2012) and this explains, why only pCON<sup>CENP-A</sup> with functional CATD were maintained in our experiments (Figure 15). Therefore, it is likely, that new CENP-A is incorporated on plasmids by HJURP.

### 5.4 Histone modifications at matured artificial centromeres

The sctetR:GFP:CENP-A fusion protein is incorporated into endogenous centromeres, if no tetO containing plasmid is present in the cells (Figure 32). We also know that sctetR:CENP-A forms nucleosomes together with H2B and without canonical H3 (Figure 34). What we do not know from our experiments is, how

sctetR:CENP-A nucleosomes are assembled at plasmid centromeres. It is unclear if tetO repeats result in a different nucleosome conformation, since these are prokaryotic sequences. Therefore, it might be that sctetR:CENP-A nucleosomes as well as canonical nucleosomes establish differently at these prokaryotic DNA structures on pCON<sup>CENP-A</sup>. However, my expectation is that sctetR:CENP-A nucleosomes at tetO repeats on plasmids are similar to endogenous CENP-A nucleosomes. It was reported, that CENP-C reshapes and stabilizes CENP-A nucleosome structure (Falk et al., 2015). Since CENP-C is also recruited to plasmids (Figure 29) it is likely that it also influences the CENP-A nucleosome structure at plasmid nucleosomes. I assume that sctetR:CENP-A nucleosomes at the tetO array are similar to endogenous centromeric nucleosomes. In order to verify this hypothesis the Cas9-dependent tandem affinity purification system, I generated for the pCON<sup>CENP-A</sup> plasmids, could be used. After plasmid purification by Cas9:TAP, mono-nucleosomes containing sctetR:CENP-A could be enriched by a tetR-IP and further analyzed according to structure and composition.

The question I addressed with the Cas9-dependent tandem affinity purification was how matured centromeres on pCON<sup>CENP-A</sup> are defined and if histone modifications are established at inheritable centromeres. It is known that at endogenous centromeres, CENP-A nucleosomes are interspersed with canonical H3 nucleosomes that show a centromere specific histone modification pattern (Fukagawa, 2017). In order to uncover histone modifications present on matured artificial centromeres on pCON<sup>CENP-A</sup>, I developed and investigated the Cas9:STREP:STREP:Flag targeting system for plasmid purification as a prerequisite for histone modification analysis by mass spectrometry. It was already shown that pull down of specific DNA by CRISPR/Cas9 and subsequent mass spectrometric analysis enables epigenome definition of a certain gene locus (Waldrip et al., 2014). We put a lot of effort into the establishment of the purification protocol. We first found, that proteins and DNA needed to be properly re-solubilized by French Press (Figure 43 and Figure 44) according to Ide et al. (Ide and Dejardin, 2015). Pull down of plasmids harboring

targeting sites was significantly increased compared to plasmids that do not contain targeting sites (Figure 41).

My initial experiment after successful pull down of Cas9 protein bound to plasmids revealed several histone modifications (Figure 46). Di-methylation of histone H3 on lysine 4 (H3K4me<sub>2</sub>), was present at artificial centromeres. It is the most prominent modification of centromeres since it is required for HJURP targeting of CENP-A to centromeres (Bergmann et al., 2011). Another important modification is mono-methylation of lysine 20 on histone H4 (H4K20me<sub>1</sub>) in CENP-A nucleosomes, which is needed for kinetochore assembly (Hori et al., 2014). However, H4K20me<sub>1</sub> was not detectable at neo-centromeres on the plasmids. In contrast, typical modifications for pericentromeric chromatin, like di- and tri-methylation of histone H3 on lysine 9 (H3K9me<sub>2/3</sub>), tri-methylation of lysine 27 on histone H3 (H3K27me<sub>3</sub>) and tri-methylation of lysine 20 on histone H4 (H4K20me<sub>3</sub>) were detectable on the plasmid system within the centromere region. This pull down with subsequent histone modification analysis was done only once because of time limitations. However, it depicted a histone modification pattern that was different from endogenous centromeres. This may have several reasons:

First, the Cas9 targeting system, I was using has off target effects. A first hint of off-target binding was already given by immune fluorescence in fixed cells (Figure 26 B), where the Cas9 protein is distributed over the whole nucleus when the reporter plasmid does not contain targeting sites. It is essential to know if Cas9 is bound to any endogenous DNA in this setting before concluding histone modifications established on the plasmid with this method. Because H3K9me<sub>3</sub> and H4K20me<sub>3</sub> are abundant histone modifications and pulling down Cas9 proteins that bind to off-target DNA, enriched in these modifications, leads to detection of these modifications in Western Blot. Therefore, the determination of possible off target effects of CRISPR/Cas9 by ChIP-seq is essential.

Second, the digest efficiency represented in Figure 42 is not exactly determined for re-solubilized and purified plasmid fragments and whole plasmids, which also

establish pericentromeric chromatin marks, were pulled down. Analyzing Cas9:TAP precipitations of whole plasmids, reveals also histone modifications that are more specific for pericentromeres.

Third possibility is that the epigenetic modification pattern of artificial centromeres is different from that of endogenous human centromeres.

Since this experiment was done only once, it is not statistically significant.

However, with the Cas9:STREP:STREP:Flag system, I generated a prerequisite for future investigations of artificial centromeres on pCON<sup>CENP-A</sup>. By tandem affinity purification of plasmids harboring inheritable centromeres that recruit minimal kinetochores, histone modifications of matured centromeres and the composition of minimal kinetochores could be uncovered. In addition, the timing of histone modification establishment during centromere maturation could be investigated with this method. And finally the structure of a whole kinetochore complex could be solved after pCON<sup>CENP-A</sup> purification by Cas9:TAP.

## 6 CONCLUSION

Human centromeres are specific regions on chromosomes that are responsible for kinetochore assembly and correct sister-chromatid segregation during mitosis. Their localization is not determined by the underlying DNA sequence. Therefore, the inheritance of centromere identity is epigenetically specified. The major hallmark for defining centromeres is the centromere specific H3 variant CENP-A. Artificial targeting of CENP-A to plasmids leads to the establishment of neo-centromeres. I used this approach to functionally investigate centromere maturation on pCON<sup>CENP-A</sup> plasmids by microscopy, cell biological and biochemical methods.

I demonstrated that artificial targeting of sctetR:CENP-A to tetO plasmids lead to the recruitment of inner and outer kinetochore components. The kinetochore at plasmid neo-centromeres distributed plasmids to daughter cells. The segregation mechanism I observed once by live cell imaging was asymmetrical, but it needs to be clarified more precise.

However, sctetR:CENP-A targeting is sufficient to establish an inheritable centromere identity on pCON<sup>CENP-A</sup> in human cells, because plasmids were stable maintained over five months. In addition, the centromere maturation on plasmids is a very fast process since only three to four cell cycles were sufficient to generate stable plasmid maintenance. Within this timeframe targeting-independent CENP-A was incorporated into artificial plasmid centromeres.

The mechanisms leading to matured and inheritable centromeres were not solved during this project. Nevertheless, I developed a CRISPR/Cas9 targeting system to purify pCON<sup>CENP-A</sup> plasmids carrying matured neo-centromeres. This tool serves as prerequisite for the analysis of epigenetic marks, like histone modifications, that establish during centromere maturation. In addition it facilitates the investigation of kinetochore complex composition and may provide an insight into kinetochore complex structure.

## 6 CONCLUSION

---

In conclusion, the pCON<sup>CENP-A</sup> plasmid is a suitable tool for investigating *de novo* centromere formation mechanisms with additional potential application in gene therapy.



**BIBLIOGRAPHY**

- Allshire, R.C., and G.H. Karpen. 2008. Epigenetic regulation of centromeric chromatin: old dogs, new tricks? *Nature reviews. Genetics*. 9:923-937.
- Amor, D.J., K. Bentley, J. Ryan, J. Perry, L. Wong, H. Slater, and K.H. Choo. 2004. Human centromere repositioning "in progress". *Proceedings of the National Academy of Sciences of the United States of America*. 101:6542-6547.
- Bailey, A.O., T. Panchenko, K.M. Sathyan, J.J. Petkowski, P.J. Pai, D.L. Bai, D.H. Russell, I.G. Macara, J. Shabanowitz, D.F. Hunt, B.E. Black, and D.R. Foltz. 2013. Posttranslational modification of CENP-A influences the conformation of centromeric chromatin. *Proceedings of the National Academy of Sciences of the United States of America*. 110:11827-11832.
- Barnhart, M.C., P.H. Kuich, M.E. Stellfox, J.A. Ward, E.A. Bassett, B.E. Black, and D.R. Foltz. 2011. HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. *J Cell Biol*. 194:229-243.
- Bassett, E.A., J. DeNizio, M.C. Barnhart-Dailey, T. Panchenko, N. Sekulic, D.J. Rogers, D.R. Foltz, and B.E. Black. 2012. HJURP uses distinct CENP-A surfaces to recognize and to stabilize CENP-A/histone H4 for centromere assembly. *Dev Cell*. 22:749-762.
- Basu, J., and H.F. Willard. 2005. Artificial and engineered chromosomes: non-integrating vectors for gene therapy. *Trends in molecular medicine*. 11:251-258.
- Bergmann, J.H., M.G. Rodriguez, N.M. Martins, H. Kimura, D.A. Kelly, H. Masumoto, V. Larionov, L.E. Jansen, and W.C. Earnshaw. 2011. Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. *Embo j*. 30:328-340.
- Bernard, P., J.-F. Maure, J.F. Partridge, S. Genier, J.-P. Javerzat, and R.C. Allshire. 2001. Requirement of Heterochromatin for Cohesion at Centromeres. *Science (New York, N.Y.)*. 294:2539-2542.
- Bjursell, G., and P. Reichard. 1973. Effects of thymidine on deoxyribonucleoside triphosphate pools and deoxyribonucleic acid synthesis in Chinese hamster ovary cells. *J Biol Chem*. 248:3904-3909.
- Black, B.E., L.E. Jansen, P.S. Maddox, D.R. Foltz, A.B. Desai, J.V. Shah, and D.W. Cleveland. 2007. Centromere identity maintained by nucleosomes assembled with histone H3 containing the CENP-A targeting domain. *Mol Cell*. 25:309-322.
- Blower, M.D., B.A. Sullivan, and G.H. Karpen. 2002. Conserved organization of centromeric chromatin in flies and humans. *Dev Cell*. 2:319-330.
- Bodor, D.L., J.F. Mata, M. Sergeev, A.F. David, K.J. Salimian, T. Panchenko, D.W. Cleveland, B.E. Black, J.V. Shah, and L.E. Jansen. 2014. The quantitative architecture of centromeric chromatin. *eLife*. 3:e02137.

- Brustel, J., N. Kirstein, F. Izard, C. Grimaud, P. Prorok, C. Cayrou, G. Schotta, A.F. Abdelsamie, J. Dejardin, M. Mechali, G. Baldacci, C. Sardet, J.C. Cadoret, A. Schepers, and E. Julien. 2017. Histone H4K20 tri-methylation at late-firing origins ensures timely heterochromatin replication. *Embo j.*
- Carroll, C.W., K.J. Milks, and A.F. Straight. 2010. Dual recognition of CENP-A nucleosomes is required for centromere assembly. *J Cell Biol.* 189:1143-1155.
- Carroll, C.W., M.C. Silva, K.M. Godek, L.E. Jansen, and A.F. Straight. 2009. Centromere assembly requires the direct recognition of CENP-A nucleosomes by CENP-N. *Nature cell biology.* 11:896-902.
- Cheeseman, I.M., and A. Desai. 2008. Molecular architecture of the kinetochore-microtubule interface. *Nature reviews. Molecular cell biology.* 9:33-46.
- Chen, B., L.A. Gilbert, B.A. Cimini, J. Schnitzbauer, W. Zhang, G.W. Li, J. Park, E.H. Blackburn, J.S. Weissman, L.S. Qi, and B. Huang. 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell.* 155:1479-1491.
- Delecluse, H.J., and W. Hammerschmidt. 2000. The genetic approach to the Epstein-Barr virus: from basic virology to gene therapy. *Molecular pathology : MP.* 53:270-279.
- Dunleavy, E.M., W. Zhang, and G.H. Karpen. 2013. Solo or doppio: how many CENP-As make a centromeric nucleosome? *Nat Struct Mol Biol.* 20:648-650.
- Dupont, C., D.R. Armant, and C.A. Brenner. 2009. Epigenetics: definition, mechanisms and clinical perspective. *Seminars in reproductive medicine.* 27:351-357.
- Earnshaw, W.C., and N. Rothfield. 1985. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma.* 91:313-321.
- Ehrhardt, A., R. Haase, A. Schepers, M.J. Deutsch, H.J. Lipps, and A. Baiker. 2008. Episomal vectors for gene therapy. *Current gene therapy.* 8:147-161.
- Falk, S.J., L.Y. Guo, N. Sekulic, E.M. Smoak, T. Mani, G.A. Logsdon, K. Gupta, L.E. Jansen, G.D. Van Duyne, S.A. Vinogradov, M.A. Lampson, and B.E. Black. 2015. Chromosomes. CENP-C reshapes and stabilizes CENP-A nucleosomes at the centromere. *Science (New York, N.Y.).* 348:699-703.
- Falk, S.J., J. Lee, N. Sekulic, M.A. Sennett, T.H. Lee, and B.E. Black. 2016. CENP-C directs a structural transition of CENP-A nucleosomes mainly through sliding of DNA gyres. *Nat Struct Mol Biol.* 23:204-208.
- Foley, E.A., and T.M. Kapoor. 2013. Microtubule attachment and spindle assembly checkpoint signalling at the kinetochore. *Nature reviews. Molecular cell biology.* 14:25-37.
- Foltz, D.R., L.E. Jansen, A.O. Bailey, J.R. Yates, 3rd, E.A. Bassett, S. Wood, B.E. Black, and D.W. Cleveland. 2009. Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. *Cell.* 137:472-484.
- Fujita, R., K. Otake, Y. Arimura, N. Horikoshi, Y. Miya, T. Shiga, A. Osakabe, H. Tachiwana, J. Ohzeki, V. Larionov, H. Masumoto, and H. Kurumizaka. 2015. Stable complex formation of CENP-B with the CENP-A nucleosome. *Nucleic acids research.* 43:4909-4922.

- Fukagawa, T. 2017. Critical histone post-translational modifications for centromere function and propagation. *Cell cycle (Georgetown, Tex.)*:1-7.
- Gascoigne, K.E., K. Takeuchi, A. Suzuki, T. Hori, T. Fukagawa, and I.M. Cheeseman. 2011. Induced ectopic kinetochore assembly bypasses the requirement for CENP-A nucleosomes. *Cell*. 145:410-422.
- Gerlich, D., T. Hirota, B. Koch, J.M. Peters, and J. Ellenberg. 2006. Condensin I stabilizes chromosomes mechanically through a dynamic interaction in live cells. *Current biology : CB*. 16:333-344.
- Guse, A., C.W. Carroll, B. Moree, C.J. Fuller, and A.F. Straight. 2011. In vitro centromere and kinetochore assembly on defined chromatin templates. *Nature*. 477:354-358.
- Hagege, H., P. Klous, C. Braem, E. Splinter, J. Dekker, G. Cathala, W. de Laat, and T. Forne. 2007. Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nature protocols*. 2:1722-1733.
- Hahn, M., S. Dambacher, S. Dulev, A.Y. Kuznetsova, S. Eck, S. Worz, D. Sadic, M. Schulte, J.P. Mallm, A. Maiser, P. Debs, H. von Melchner, H. Leonhardt, L. Schermelleh, K. Rohr, K. Rippe, Z. Storchova, and G. Schotta. 2013. Suv4-20h2 mediates chromatin compaction and is important for cohesin recruitment to heterochromatin. *Genes Dev*. 27:859-872.
- Hasson, D., T. Panchenko, K.J. Salimian, M.U. Salman, N. Sekulic, A. Alonso, P.E. Warburton, and B.E. Black. 2013. The octamer is the major form of CENP-A nucleosomes at human centromeres. *Nat Struct Mol Biol*. 20:687-695.
- Henikoff, S., T. Furuyama, and K. Ahmad. 2004. Histone variants, nucleosome assembly and epigenetic inheritance. *Trends in Genetics*. 20:320-326.
- Heun, P., S. Erhardt, M.D. Blower, S. Weiss, A.D. Skora, and G.H. Karpen. 2006. Mislocalization of the *Drosophila* centromere-specific histone CID promotes formation of functional ectopic kinetochores. *Dev Cell*. 10:303-315.
- Hirt, B. 1966. Evidence for semiconservative replication of circular polyoma DNA. *Proceedings of the National Academy of Sciences of the United States of America*. 55:997-1004.
- Hoffmann, S., M. Dumont, V. Barra, P. Ly, Y. Nechemia-Arbely, M.A. McMahon, S. Herve, D.W. Cleveland, and D. Fachinetti. 2016. CENP-A Is Dispensable for Mitotic Centromere Function after Initial Centromere/Kinetochore Assembly. *Cell reports*. 17:2394-2404.
- Hori, T., W.H. Shang, A. Toyoda, S. Misu, N. Monma, K. Ikeo, O. Molina, G. Vargiu, A. Fujiyama, H. Kimura, W.C. Earnshaw, and T. Fukagawa. 2014. Histone H4 Lys 20 monomethylation of the CENP-A nucleosome is essential for kinetochore assembly. *Dev Cell*. 29:740-749.
- Hu, H., Y. Liu, M. Wang, J. Fang, H. Huang, N. Yang, Y. Li, J. Wang, X. Yao, Y. Shi, G. Li, and R.M. Xu. 2011. Structure of a CENP-A-histone H4 heterodimer in complex with chaperone HJURP. *Genes Dev*. 25:901-906.
- Ide, S., and J. Dejardin. 2015. End-targeting proteomics of isolated chromatin segments of a mammalian ribosomal RNA gene promoter. *Nature communications*. 6:6674.

- Jansen, L.E., B.E. Black, D.R. Foltz, and D.W. Cleveland. 2007. Propagation of centromeric chromatin requires exit from mitosis. *J Cell Biol.* 176:795-805.
- Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J.A. Doudna, and E. Charpentier. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N.Y.)*. 337:816-821.
- Kato, H., J. Jiang, B.R. Zhou, M. Rozendaal, H. Feng, R. Ghirlando, T.S. Xiao, A.F. Straight, and Y. Bai. 2013. A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science (New York, N.Y.)*. 340:1110-1113.
- Kim, S., and H. Yu. 2015. Multiple assembly mechanisms anchor the KMN spindle checkpoint platform at human mitotic kinetochores. *J Cell Biol.* 208:181-196.
- Kirchmaier, A.L., and B. Sugden. 1995. Plasmid maintenance of derivatives of oriP of Epstein-Barr virus. *Journal of virology*. 69:1280-1283.
- Kornberg, R.D. 1974. Chromatin structure: a repeating unit of histones and DNA. *Science (New York, N.Y.)*. 184:868-871.
- Kornberg, R.D., and J.O. Thomas. 1974. Chromatin structure; oligomers of the histones. *Science (New York, N.Y.)*. 184:865-868.
- Logsdon, G.A., E.J. Barrey, E.A. Bassett, J.E. DeNizio, L.Y. Guo, T. Panchenko, J.M. Dawicki-McKenna, P. Heun, and B.E. Black. 2015. Both tails and the centromere targeting domain of CENP-A are required for centromere establishment. *J Cell Biol.* 208:521-531.
- Lufino, M.M., P.A. Edser, and R. Wade-Martins. 2008. Advances in high-capacity extrachromosomal vector technology: episomal maintenance, vector delivery, and transgene expression. *Molecular therapy : the journal of the American Society of Gene Therapy*. 16:1525-1538.
- Luger, K., A.W. Mader, R.K. Richmond, D.F. Sargent, and T.J. Richmond. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*. 389:251-260.
- Ma, H., A. Naseri, P. Reyes-Gutierrez, S.A. Wolfe, S. Zhang, and T. Pederson. 2015. Multicolor CRISPR labeling of chromosomal loci in human cells. *Proceedings of the National Academy of Sciences of the United States of America*. 112:3002-3007.
- Marshall, O.J., A.C. Chueh, L.H. Wong, and K.H. Choo. 2008. Neocentromeres: new insights into centromere structure, disease development, and karyotype evolution. *American journal of human genetics*. 82:261-282.
- Masumoto, H., H. Masukata, Y. Muro, N. Nozaki, and T. Okazaki. 1989. A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *J Cell Biol.* 109:1963-1973.
- Masumoto, H., M. Nakano, and J. Ohzeki. 2004. The role of CENP-B and alpha-satellite DNA: de novo assembly and epigenetic maintenance of human centromeres. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology*. 12:543-556.
- McClintock, B. 1941. The Stability of Broken Ends of Chromosomes in Zea Mays. *Genetics*. 26:234-282.

- McEwen, B.F., Y. Dong, and K.J. VandenBeldt. 2007. Using electron microscopy to understand functional mechanisms of chromosome alignment on the mitotic spindle. *Methods in cell biology*. 79:259-293.
- McKinley, K.L., and I.M. Cheeseman. 2016. The molecular basis for centromere identity and function. *Nature reviews. Molecular cell biology*. 17:16-29.
- Mendiburo, M.J., J. Padeken, S. Fulop, A. Schepers, and P. Heun. 2011. Drosophila CENH3 is sufficient for centromere formation. *Science (New York, N.Y.)*. 334:686-690.
- Middleton, T., and B. Sugden. 1992. EBNA1 can link the enhancer element to the initiator element of the Epstein-Barr virus plasmid origin of DNA replication. *Journal of virology*. 66:489-495.
- Mullis, K., F. Faloona, S. Scharf, R. Saiki, G. Horn, and H. Erlich. 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor symposia on quantitative biology*. 51 Pt 1:263-273.
- Musacchio, A., and A. Desai. 2017. A Molecular View of Kinetochore Assembly and Function. *Biology*. 6.
- Nakano, M., S. Cardinale, V.N. Noskov, R. Gassmann, P. Vagnarelli, S. Kandels-Lewis, V. Larionov, W.C. Earnshaw, and H. Masumoto. 2008. Inactivation of a human kinetochore by specific targeting of chromatin modifiers. *Dev Cell*. 14:507-522.
- Nakashima, H., M. Nakano, R. Ohnishi, Y. Hiraoka, Y. Kaneda, A. Sugino, and H. Masumoto. 2005. Assembly of additional heterochromatin distinct from centromere-kinetochore chromatin is required for de novo formation of human artificial chromosome. *J Cell Sci*. 118:5885-5898.
- Nishino, T., K. Takeuchi, K.E. Gascoigne, A. Suzuki, T. Hori, T. Oyama, K. Morikawa, I.M. Cheeseman, and T. Fukagawa. 2012. CENP-T-W-S-X forms a unique centromeric chromatin structure with a histone-like fold. *Cell*. 148:487-501.
- Ohzeki, J., J.H. Bergmann, N. Kouprina, V.N. Noskov, M. Nakano, H. Kimura, W.C. Earnshaw, V. Larionov, and H. Masumoto. 2012. Breaking the HAC Barrier: histone H3K9 acetyl/methyl balance regulates CENP-A assembly. *Embo j*. 31:2391-2402.
- Ohzeki, J., M. Nakano, T. Okada, and H. Masumoto. 2002. CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA. *J Cell Biol*. 159:765-775.
- Okada, T., J. Ohzeki, M. Nakano, K. Yoda, W.R. Brinkley, V. Larionov, and H. Masumoto. 2007. CENP-B controls centromere formation depending on the chromatin context. *Cell*. 131:1287-1300.
- Okamoto, Y., M. Nakano, J. Ohzeki, V. Larionov, and H. Masumoto. 2007. A minimal CENP-A core is required for nucleation and maintenance of a functional human centromere. *Embo j*. 26:1279-1291.
- Olins, A.L., and D.E. Olins. 1974. Spheroid chromatin units (v bodies). *Science (New York, N.Y.)*. 183:330-332.
- Palmer, D.K., and R.L. Margolis. 1985. Kinetochore components recognized by human autoantibodies are present on mononucleosomes. *Mol Cell Biol*. 5:173-186.

- Peters, A.H., S. Kubicek, K. Mechtler, R.J. O'Sullivan, A.A. Derijck, L. Perez-Burgos, A. Kohlmaier, S. Opravil, M. Tachibana, Y. Shinkai, J.H. Martens, and T. Jenuwein. 2003. Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol Cell*. 12:1577-1589.
- Pich, D., S. Humme, M.P. Spindler, A. Schepers, and W. Hammerschmidt. 2008. Conditional gene vectors regulated in cis. *Nucleic acids research*. 36:e83.
- Ribeiro, S.A., J.C. Gatlin, Y. Dong, A. Joglekar, L. Cameron, D.F. Hudson, C.J. Farr, B.F. McEwen, E.D. Salmon, W.C. Earnshaw, and P. Vagnarelli. 2009. Condensin regulates the stiffness of vertebrate centromeres. *Molecular biology of the cell*. 20:2371-2380.
- Ribeiro, S.A., P. Vagnarelli, Y. Dong, T. Hori, B.F. McEwen, T. Fukagawa, C. Flors, and W.C. Earnshaw. 2010. A super-resolution map of the vertebrate kinetochore. *Proceedings of the National Academy of Sciences of the United States of America*. 107:10484-10489.
- Riu, E., Z.Y. Chen, H. Xu, C.Y. He, and M.A. Kay. 2007. Histone modifications are associated with the persistence or silencing of vector-mediated transgene expression in vivo. *Molecular therapy : the journal of the American Society of Gene Therapy*. 15:1348-1355.
- Ross, J.E., K.S. Woodlief, and B.A. Sullivan. 2016. Inheritance of the CENP-A chromatin domain is spatially and temporally constrained at human centromeres. *Epigenetics Chromatin*. 9:20.
- Sambrook, J., and D.W. Russell. 2006. The inoue method for preparation and transformation of competent e. Coli: "ultra-competent" cells. *CSH protocols*. 2006.
- Sander, J.D., and J.K. Joung. 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature biotechnology*. 32:347-355.
- Schalch, T., and F.A. Steiner. 2016. Structure of centromere chromatin: from nucleosome to chromosomal architecture. *Chromosoma*.
- Schepers, A., M. Ritzi, K. Bousset, E. Kremmer, J.L. Yates, J. Harwood, J.F. Diffley, and W. Hammerschmidt. 2001. Human origin recognition complex binds to the region of the latent origin of DNA replication of Epstein-Barr virus. *Embo j*. 20:4588-4602.
- Scott, K.C., and B.A. Sullivan. 2014. Neocentromeres: a place for everything and everything in its place. *Trends in genetics : TIG*. 30:66-74.
- Sekulic, N., E.A. Bassett, D.J. Rogers, and B.E. Black. 2010. The structure of (CENP-A-H<sub>4</sub>)(<sub>2</sub>) reveals physical features that mark centromeres. *Nature*. 467:347-351.
- Sekulic, N., and B.E. Black. 2012. Molecular underpinnings of centromere identity and maintenance. *Trends in biochemical sciences*. 37:220-229.
- Sims, R.J., 3rd, K. Nishioka, and D. Reinberg. 2003. Histone lysine methylation: a signature for chromatin function. *Trends in genetics : TIG*. 19:629-639.
- Steiner, F.A., and S. Henikoff. 2014. Holocentromeres are dispersed point centromeres localized at transcription factor hotspots. *eLife*. 3:e02025.
- Steiner, F.A., and S. Henikoff. 2015. Diversity in the organization of centromeric chromatin. *Current Opinion in Genetics & Development*. 31:28-35.

- Struhl, K. 1998. Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev.* 12:599-606.
- Sullivan, B.A., and G.H. Karpen. 2004. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat Struct Mol Biol.* 11:1076-1083.
- Sullivan, K.F., M. Hechenberger, and K. Masri. 1994. Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. *J Cell Biol.* 127:581-592.
- Tachiwana, H., W. Kagawa, T. Shiga, A. Osakabe, Y. Miya, K. Saito, Y. Hayashi-Takanaka, T. Oda, M. Sato, S.Y. Park, H. Kimura, and H. Kurumizaka. 2011. Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature.* 476:232-235.
- Thomae, A.W., J. Baltin, D. Pich, M.J. Deutsch, M. Ravasz, K. Zeller, M. Gossen, W. Hammerschmidt, and A. Schepers. 2011. Different roles of the human Orc6 protein in the replication initiation process. *Cellular and molecular life sciences : CMLS.* 68:3741-3756.
- Thomae, A.W., D. Pich, J. Brocher, M.P. Spindler, C. Berens, R. Hock, W. Hammerschmidt, and A. Schepers. 2008. Interaction between HMGA1a and the origin recognition complex creates site-specific replication origins. *Proceedings of the National Academy of Sciences of the United States of America.* 105:1692-1697.
- Tomkiel, J., C.A. Cooke, H. Saitoh, R.L. Bernat, and W.C. Earnshaw. 1994. CENP-C is required for maintaining proper kinetochore size and for a timely transition to anaphase. *J Cell Biol.* 125:531-545.
- Waddington, C.H. 1968. Towards a theoretical biology. *Nature.* 218:525-527.
- Waldrip, Z.J., S.D. Byrum, A.J. Storey, J. Gao, A.K. Byrd, S.G. Mackintosh, W.P. Wahls, S.D. Taverna, K.D. Raney, and A.J. Tackett. 2014. A CRISPR-based approach for proteomic analysis of a single genomic locus. *Epigenetics.* 9:1207-1211.
- Warburton, P.E., C.A. Cooke, S. Bourassa, O. Vafa, B.A. Sullivan, G. Stetten, G. Gimelli, D. Warburton, C. Tyler-Smith, K.F. Sullivan, G.G. Poirier, and W.C. Earnshaw. 1997. Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. *Current biology : CB.* 7:901-904.
- Wei, R.R., J. Al-Bassam, and S.C. Harrison. 2007. The Ndc80/HEC1 complex is a contact point for kinetochore-microtubule attachment. *Nat Struct Mol Biol.* 14:54-59.
- Westhorpe, F.G., C.J. Fuller, and A.F. Straight. 2015. A cell-free CENP-A assembly system defines the chromatin requirements for centromere maintenance. *J Cell Biol.* 209:789-801.
- Wu, C., and J.R. Morris. 2001. Genes, genetics, and epigenetics: a correspondence. *Science (New York, N.Y.).* 293:1103-1105.
- Yates, J.L., S.M. Camiolo, and J.M. Bashaw. 2000. The minimal replicator of Epstein-Barr virus oriP. *Journal of virology.* 74:4512-4522.
- Yuan, G., and B. Zhu. 2012. Histone variants and epigenetic inheritance. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms.* 1819:222-229.

## APPENDIX

### Establishment of a suitable tetR antibody for ChIP and Co-IP

First, for the functional and biochemical characterization of the artificial centromere that is established after initial targeting of CENP-A to the plasmids, it is important to show, that the targeting of sctetR:CENP-A itself is functional. An antibody is essential to precipitate the sctetR:CENP-A protein and to analyze the co-precipitating DNA fragments in a ChIP experiment. The commercial tetR antibodies, which were tested for ChIP in our laboratory before this work, are not of sufficient quality for the planned experiments. Therefore, I decided to establish new monoclonal tetR antibodies in close collaboration with the Monoclonal Antibody Core Facility of the Helmholtz Centre Munich.

Different tetR peptides, tetR1, tetR2 and tetR, were used for the immunization of rats and mice. After fusion splenic B-cells of the immunized animals with a myeloblastoma cell line, primary supernatants containing immune globulins were collected to test the generation of specific antibodies. The core facility already determined the immune globulin class and the subtype of candidate antibodies.

In the further experiments the specificity and the suitability of these antibody supernatants to work in ChIP were described.



## **tetR antibody related methods**

### **tetR antibody validation**

New monoclonal  $\alpha$ -tetR antibodies were established in collaboration with the Monoclonal Antibody Core Facility of the Helmholtz Centre Munich. I received the primary antibody supernatants and determined their specificity in recognizing tetR fusion proteins by Western Blot (chapter 3.2.5). Positive clones were then also analyzed in their immune precipitation capacity by performing an immune precipitation according to the ChIP protocol (chapter 3.2.7) and visualizing precipitated tetR proteins by Western Blot. Cells producing primary antibodies with high immune precipitation capacity were then used for generation of stable hybridoma cells by the core facility. Supernatants of these stable clones were then tested again according to their subtype and specificity. In addition immune precipitation and ChIP efficiency was determined again.

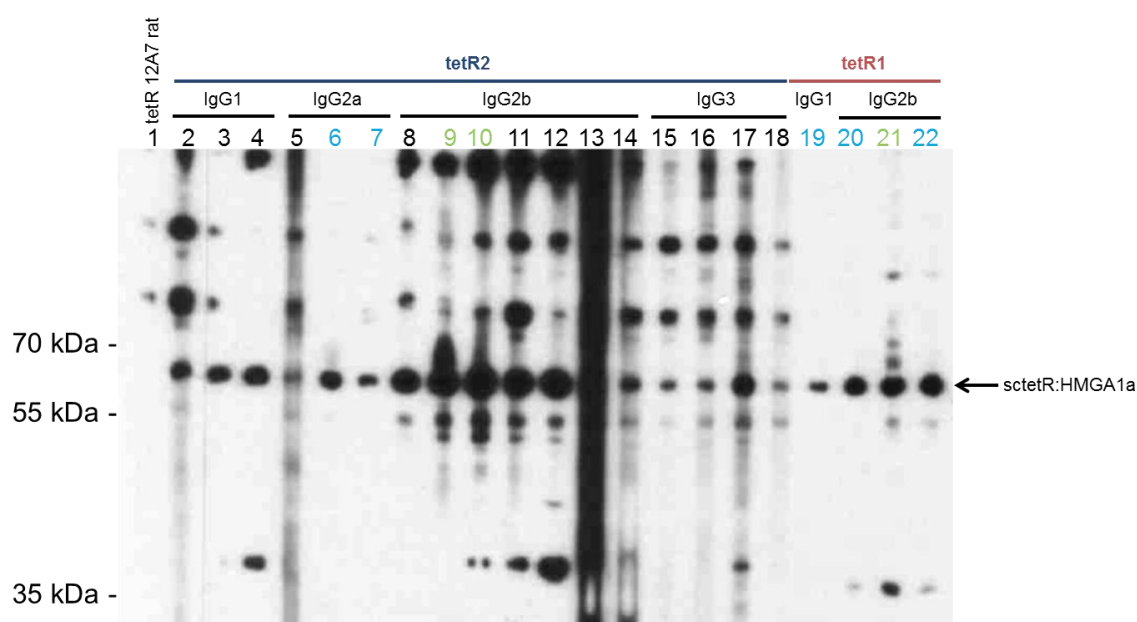
### **Covalent coupling of antibodies to protein G sepharose beads**

For co-immune precipitation experiments the tetR antibody was covalently coupled to sepharose beads to enhance IP efficiency. Since the tetR antibody used for co-immune precipitation was the 31B3 mouse-IgG2b  $\alpha$ -tetR1, protein G coated sepharose beads were utilized.

In order to couple antibodies to protein G beads, 500  $\mu$ l beads were washed first 3x with 10 ml PBS. For washing beads, these were resuspended in the respective buffer and centrifuged at 1000 g for 5 min at 4 °C. After washing, beads were resuspended in 10 ml hybridoma supernatant of 31B3  $\alpha$ -tetR antibody and incubated for 1 h at room temperature on the roller. Beads were washed 2x in 5 ml Sodium-borate buffer. Coupling of antibody to beads was performed by adding 20 mM DMP containing Sodium-borate buffer and incubation for 30 min at room temperature on the roller. Beads were washed once with 5 ml Ethanolamine buffer and afterwards incubated for 2 h in Ethanolamine buffer at room temperature. After washing beads again 2x with 5 ml PBS, beads were resuspended in 500  $\mu$ l PBS+0.02 % Sodium azide and stored at 4 ° until use.

### Screening of primary antibody supernatants

The primary supernatants, which I obtained from the antibody core facility, were first tested according to their specificity to recognize a sctetR fusion protein. Therefore cell lysates of cells expressing a sctetR:HMGA1a fusion protein were analyzed by Western Blot. In general the primary antibodies originating from mouse cells were more sensitive in detecting the specific sctetR:HMGA1a signal and had lower background signals than those obtained from rats (data not shown). For that reason I concentrated on mouse antibodies for further validation.



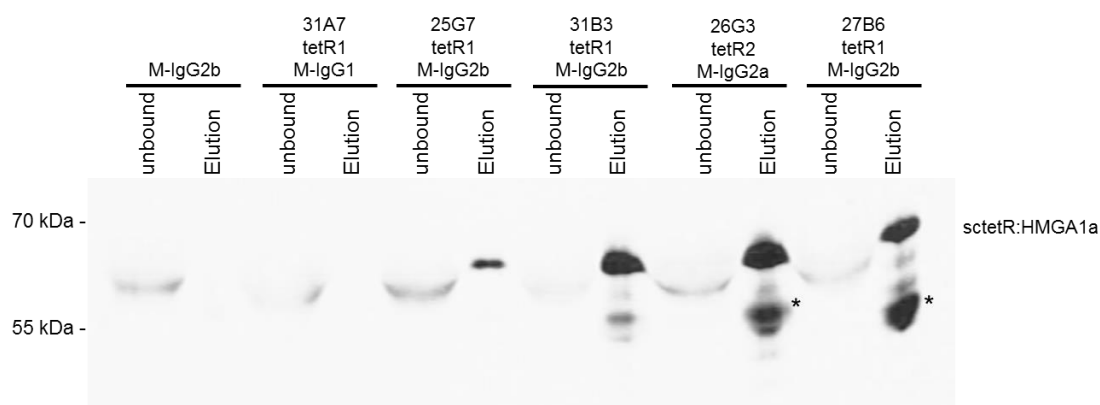
**Appendix Figure 1: Western Blot of sctetR:HMGA1a<sup>+</sup> cell lysate with different tetR antibodies**

The predicted molecular weight of sctetR:HMGA1a is 65 kDa. In a Western Blot the  $\alpha$ -HMGA1 antibody shows a signal between 55 and 70 kDa. All primary antibody supernatants depicted here show a specific signal for sctetR:HMGA1a. The difference between the tested antibody supernatants is in the background signals appearing at higher molecular weights. Blue lines 6, 7, 19, 20 and 22 display no additional unwanted signal, these are very specific antibodies. Green lines 9, 10 and 21 also show specific recognition, but also weak background. Black lines either have too high background recognition or an immune globulin subtype, not optimal for protein A or G purification in ChIP and IP. (n=1)

In Appendix Figure 1 a Western Blot of different primary mouse antibody supernatants is shown. All mouse antibodies recognized the sctetR:HMGA1a fusion protein with a molecular weight of 65 kDa. For the generation of these antibodies two different peptides, tetR1 and tetR2, were initially used for immunization. The lanes in this blot were arranged according to the original peptide and the immune globulin subclass of the resulting antibodies. For immune precipitation experiments IgG2 subclass antibodies are the most promising, because they show a high affinity to protein G that was used for precipitation of the antibody-target protein complex. Therefore stable hybridoma cells of antibodies from lanes 6 (26G3), 7 (30B11), 20 (25G7) and 22 (31B3) (Appendix Figure 1; blue marked lanes) were established by the antibody core facility. The tetR1 IgG1 antibody (31A7) (Appendix Figure 1; lane 19) was also used for generation of stable hybridoma cells because of its high specificity. The antibodies marked in green (Appendix Figure 1; lanes 9 (26H4), 10 (28B10) and 21 (27B6)) were kept as backup antibodies for the case that no stable cells could be generated out of the other clones.

### **Immune precipitation with primary antibody supernatants**

For the further validation of the tetR antibodies immune precipitations of cell extract generated of cells and cross-linked for 5 min with 1 % formaldehyde were performed. The immune precipitates of different primary antibodies was determined by Western Blot of the sctetR:HMGA1a fusion protein by comparing the unbound and eluted fraction.



### Appendix Figure 2: Immune precipitation with different primary antibody supernatants

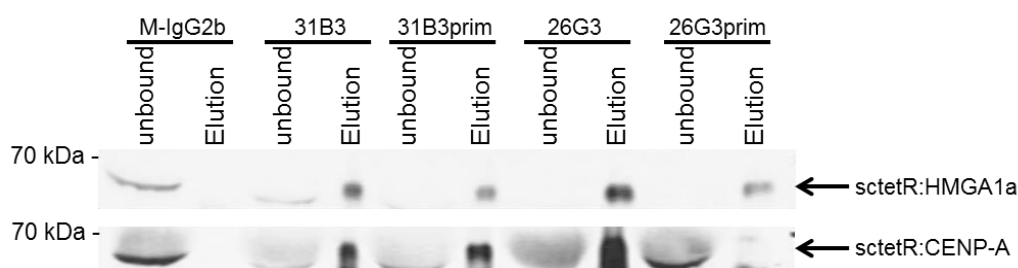
An immune precipitation (IP) with different primary mouse (M-IgG) tetR antibody supernatants was performed according to the ChIP protocol (cross-linked and sonicated). 31B3 and 26G3 give a high IP efficiency (low unbound, strong elution), 31A7 does not work in IP and 25G7 only shows a weaker elution than 31B3, 26G3 and 27B6. 27B6 is also suitable for IP, but in this IP a high degradation of sctetR:HMGA1a was observed. (n=1)

This Western Blot analysis (Appendix Figure 2) reveals that the different antibodies show huge differences in their IP efficiency of their target protein. The tetR1, M-IgG1 antibody 31A7 did not precipitate at all. The efficiency of 25G7 (tetR1, M-IgG2b) was also low compared to the three other antibodies 31B3, 26G3 and 27B6. The strong second band appearing for the antibodies 26G3 and 27B6 (Appendix Figure 2; \*) might be the sctetR degradation product of sctetR:HMGA1a, since sctetR without HMGA1a fusion has a molecular weight of 55 kDa.

According to the specificity shown in Appendix Figure 1 and the IP efficiency depicted in Appendix Figure 2 the generation of stable hybridoma cells was continued for the clones 31B3 and 26G3 in the antibody core facility.

### ChIP with primary vs. stable generated antibody supernatant

The next step of the tetR antibody validation was to compare the original primary antibody supernatant with the supernatant obtained from stable hybridoma clones of the antibodies 31B3 and 26G3 in their IP (Appendix Figure 3) and ChIP efficiency.



**Appendix Figure 3: Immune precipitation with 31B3 and 26G3 primary and stable supernatant**

An IP with the primary and stable (supernatant obtained from hybridoma cell clones) tetR antibody supernatants was performed according to the ChIP protocol. IP efficiencies are depicted by Western Blot with the rat  $\alpha$ -tetR antibody 12A7. The highest IP efficiency is observed with the stable 26G3 antibody. (n=1)

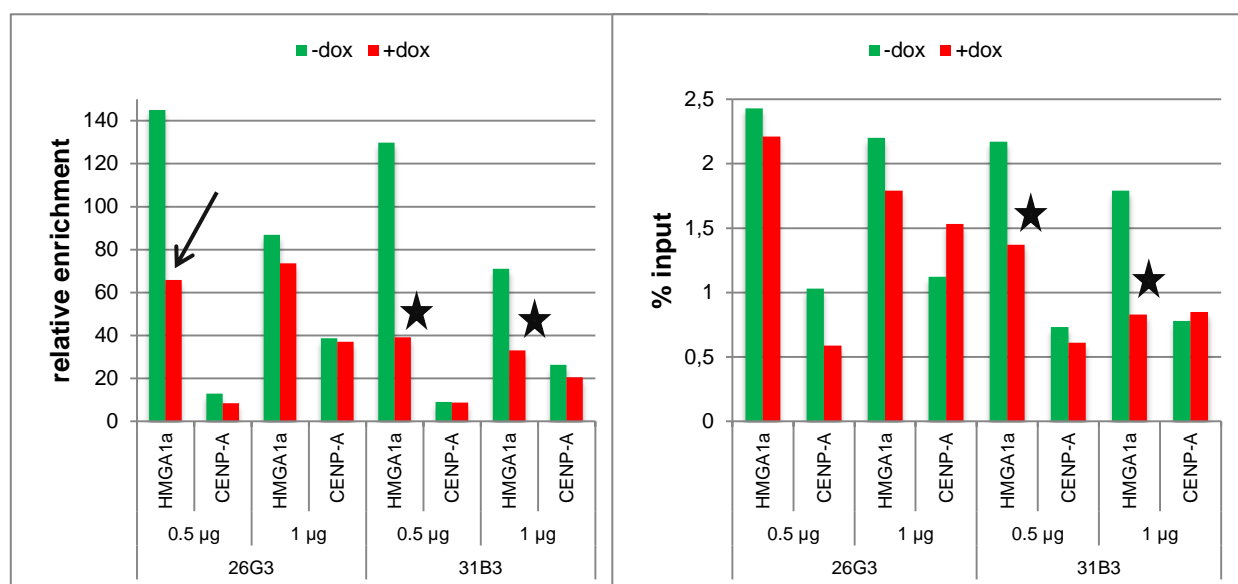
The highest IP efficiency, according to the Western Blot signal, was obtained by the 26G3 antibody supernatant of hybridoma cells. But also the 31B3 antibody shows an efficient and specific binding to the sctetR target proteins. The signals obtained from the primary supernatants in this experiment were slightly lower (Appendix Figure 3).

The same samples used for the Western Blot, were also analyzed by qPCR to determine the ChIP capacity for these two antibodies.

As qPCR readout of the tetR ChIP different regions on the plasmid were chosen. The primer pairs covered the region in close proximity of the tetO repeats (tetO), the DS primer is next to DS element and the reference primer pair is several kilobases away from DS and tetO (Figure 20 A).

Since the readout of the ChIP of the samples was not clear (data not shown) and the tetR antibodies need to detect slight differences in binding capacity of the tetR proteins to tetO sites, a second analysis was performed.

HEK293EBNA1<sup>+</sup> cells transfected with 0.5  $\mu$ g or 1  $\mu$ g the tetO-DS reporter plasmids, expressing either sctetR:HMGA1a or sctetR:CENP-A, were incubated with 2  $\mu$ g/ml doxycycline for 24 h. Doxycycline inhibits the binding of sctetR to the tetO sites on the reporter plasmids. After ChIP with the tetR antibody it is expected that a difference in binding of tetR to tetO is detectable, at least for the sctetR:HMGA1a targeting (Pich et al., 2008).



**Appendix Figure 4: ChIP with and without doxycycline**

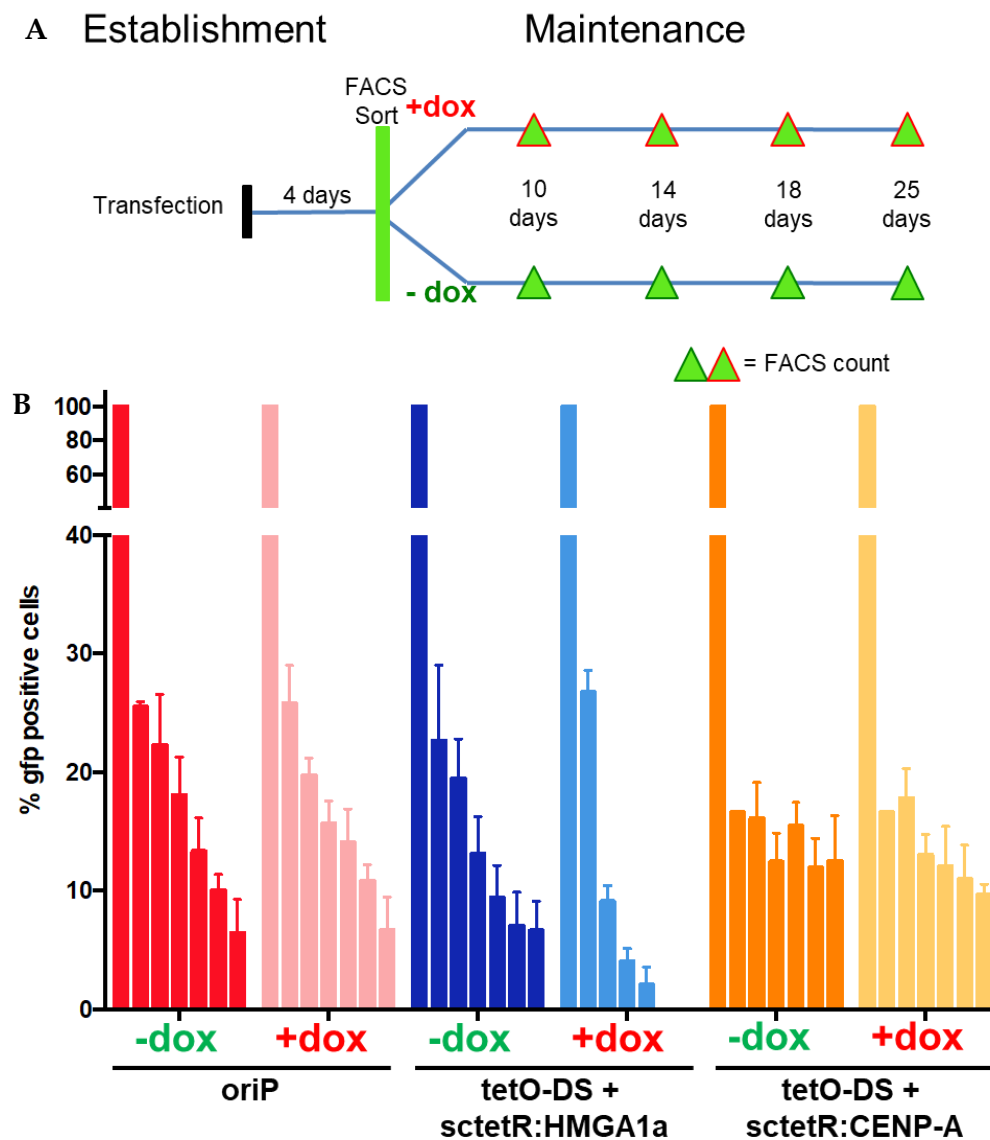
Both graphs represent the same ChIP experiment. The relative enrichment against a mouse IgG control IP is shown on the left and the % input values are depicted on the right. Both representations demonstrate that 31B3 is more sensitive to doxycycline treatment (black stars). The qPCR product is amplified by the tetO primer pair next to tetO sites. (n=1)

By comparing the sensitivities of the 26G3 and 31B3 antibodies it is striking that only the 31B3 antibody detects a difference in binding of sctetR:HMGA1a to tetO (0.5  $\mu$ g and 1  $\mu$ g plasmid transfected) and with both representations of the qPCR readout (relative enrichment and % input) (Appendix Figure 4; black stars). The 26G3 only depicts 50 % reduced binding in the relative enrichment representation when 0.5  $\mu$ g plasmid was transfected (Appendix Figure 4; black arrow).

In conclusion, the 31B3  $\alpha$ -tetR antibody is more sensitive in detecting the targeting differences upon doxycycline treatment and shows high specificity in tetR protein recognition. Consequently, this antibody was used for further experiments.

### Targeting independent plasmid maintenance after 4 days

HEK293EBNA1<sup>+</sup> cells were transfected with the different reporter plasmids (*oriP*, *sctetR:HMGA1a* *tetO-DS* and *pCON<sup>CENP-A</sup>*) and incubated for four days. After this incubation time (= establishment phase) cells were FACS sorted according to *gfp* expression and performed as already described in 4.6.2.



**Appendix Figure 5: Targeting independent plasmid maintenance 4 days after establishment**

A) Timeline of experimental setup. 4 days after transfection cells are sorted according to *gfp* expression and then splitted into condition with and without doxycycline. Plasmid maintenance was analyzed by FACS count every 4 to 6 days.

B) Plasmid maintenance after 4 days establishment time. (n=3)



## ACKNOWLEDGEMENT

First of all I would like to thank Dr. Aloys Schepers for his awesome supervision and fruitful discussions during my PhD. He always motivated me to follow my ambitions and supported me in supervising master students. Furthermore, he facilitated my contribution to many conferences and he always had an open door to listen to my problems and wishes.

I am very grateful to our collaboration partner Dr. Patrick Heun, not only for useful instructions into immune fluorescence during my lab visit, but also for his contribution and discussions as a member of my thesis advisory committee.

Many thanks also to Dr. Axel Imhof for his suggestions as member of my thesis advisory committee and for his help in setting up the tandem affinity purification protocol.

For the collaboration and discussions in establishing and performing live cell experiments, I would like to thank Dr. Steffen Dietzel and Dr. Andreas Thomae.

I am also very thankful to the people helping me with using the French Press, Ina Zimmer, technician in the institute of biochemical plant pathology of the Helmholtz Zentrum München and Katja Katzer, PhD student in the group of Dr. Parniske from the LMU München.

Special thank goes to Alejandro Freyermuth, the master student performing the setting up of the Cas9 purification protocol under my supervision.

Finally, I also thank all (past) group members, especially Kris and Petra for their contribution to my experiments, Lara for her collaboration in the plasmid maintenance approach and the Schepers Racing Team (Nina, Julia and Sarah) for an awesome working atmosphere, scientific and less scientific discussions, exchange of (Thermomix) recipes and song challenges.

---